DOCUMENT RESUME

ED 350 006 IR 054 190

AUTHOR	Cleveland, Gary
TITLE	Research Networks and Libraries. Applications and Issues for a Global Information Network. UDT Series on Data Communication Technologies and Standards for Libraries, Report #1.
INSTITUTION	International Federation of Library Associations and Institutions, Ottawa (Ontario). International Office for Universal Dataflow & Telecommunications.
REPORT NO	ISBN-0-9694214-1-9; ISSN-1018-0311
PUB DATE	91
NOTE	230p.; For related documents, see IR 054 175, IR 054 178-179, IR 054 182-183, and IR 054 197.
PUB TYPE	Information Analyses (070) -- Reports - Descriptive (141)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Computer Networks; Developing Nations; Foreign Countries; Futures (of Society); Information Dissemination; *Information Networks; International Programs; *Library Automation; *Library Services; Policy Formation; Research and Development; *Standards
IDENTIFIERS	Information Infrastructure; *Protocols (Networking); *Research Networks

ABSTRACT

This paper is intended to provide interested librarians with a basic knowledge of research networks (as opposed to library networks), focusing specifically on the manner in which they can be used by the international library community to support library functions and improve access to bibliographic information, as well as the technical and policy issues involved in research network use. (Research networks are defined as networks developed within the academic community to provide the infrastructure for a variety of networking tasks, while library networks are designed specifically to support the nuts and bolts operations of libraries.) While major projects involving library/research network connectivity are discussed, the primary focus of the study is the potential for research network use and available standards. After an introductory initial chapter, the second chapter provides an overview of network concepts, including network types and protocols. Library applications of research networks are addressed in the third chapter, including applications for bibliographic provision, document supply, and technical services. In the fourth chapter, library and research network projects in North America, Europe, Australia, and Japan are described. Technical and policy issues, including governance, intellectual property rights, appropriate use, access, and staffing, are considered in the fifth chapter. Chapter 5 also discusses the outlook for developing nations; chapter 6 looks briefly at future network initiatives; and chapter 7 summarizes benefits of and obstacles to library use of research networks. Relevant standards are listed; and the appendix provides background information (i.e., name, date started, purpose, location, configuration, governance, financing, membership, and connections) for networks in North America, Europe, Eastern Europe, Australasia, South East Asia, Southern Asia, the Far East, Latin America and the Caribbean, the Middle East, and Africa. (MES)

ED350006

International Federation of Library Associations and Institutions

Universal Dataflow and Telecommunications

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Research Networks and Libraries

Applications and Issues for a Global Information Network

Gary Cleveland

IR 054190

1991

UDT Series on Data Communication Technologies and Standards for Libraries

Report # 1

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
Paula Tallim
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

International Federation of Library Associations and Institutions

Universal Dataflow and Telecommunications

Research Networks and Libraries

Applications and Issues for a Global Information Network

Gary Cleveland

1991

UDT Series on Data Communication Technologies and Standards for Libraries

Report # 1

UDT Series on Data Communication Technologies and Standards for Libraries - Report #1

IFLA International Office for Universal Dataflow and Telecommunications
c/o National Library of Canada
395 Wellington Street
Ottawa, CANADA K1A 0N4

This publication is printed on permanent paper.

Canadian Cataloguing in Publication Data

Cleveland, Gary, 1958-

Research networks and libraries

(UDT series on data communication technologies and standards for libraries, ISSN 1018-0311 ; report #1)
Includes bibliographical references.
ISBN 0-9694214-1-9

1. Information networks. 2. Library information networks. 3. Information retrieval. I. International Federation of Library Associations and Institutions. Universal Dataflow and Telecommunications Core Program II. Title. III. Series.

Z674.4.C54 1991 025.04 C91-090312-3

CONTENTS

Foreword xi
Acknowledgements xii

1. Introduction 1

Research networks and libraries 1
Goals of the paper 2
Research networks versus library networks 2
Organization of the paper 4

2. Overview of Network Concepts 7

2.1 Network types 9
2.1.1 Local Area Networks (LANs) 9
2.1.2 Metropolitan Area Networks (MANs) 11
2.1.3 Wide Area Networks (WANs) 11
2.1.3.1 Circuit-switched networks 12
2.1.3.2 Packet-switched networks 12
2.1.4 Public Data Networks (PDNs) 13
2.1.5 Value-added networks 14
2.1.6 Integrated Services Digital Network (ISDN) 14
2.1.7 Store-and-Forward Networks 15
2.1.8 Internetworks 16

2.2 Overview of protocols 16
2.2.1 Layering, transparency, and virtual facilities 17
Layering 17
Layer transparency 18
Layer envelopes 18
Layer independence 18
Advantages and disadvantages of layering 19
2.2.2 Modes of Communication 19
2.2.3 The OSI Reference Model 20

2.2.4 Major protocol suites 22
2.2.4.1 The OSI Protocol Suite 22
2.2.4.2 TCP/IP 24
2.2.4.3 DECNET 25
2.2.4.4 Coloured Book 25
2.2.4.5 SNA 25
2.2.4.6 Other notable protocols 26
2.2.5 Protocols for Specific Applications 26
2.2.5.1 Electronic mail 27
SMTP 27
BITNET BSMTP 28
Grey Book 28
X.400 28
2.2.5.2 Remote login 29
TELNET 29
rlogin 30
Triple-X 30
Green Book 30
Virtual Terminal 30
X Window and Motif 31
2.2.5.3 File Transfer 31
FTP 31
Blue Book 32
FTAM 32
2.2.5.4 Remote Job Entry 32
2.2.5.5 Directory Services 33
DNS 33
JANET DNS 35
X.500 Directory service 35
2.2.5.6 OSI library-oriented protocols 37
Z39.50 and SR 37
The ILL Protocol 39

3. Library Applications of Research Networks 41

3.1 Applications for bibliographic information provision 42
3.1.1 Information retrieval 42
Use of the research network infrastructure 44
Distributed information retrieval 45
The Z39.50 protocol 46
Employment 46
3.1.2 Selective Dissemination of Information (SDI) 46

3.2 Applications for document supply 47
3.2.1 Interlibrary loan 47
Use of the research network infrastructure 48
The ILL protocol 49
Message handling systems and ILL 50
Employment 50
3.2.2 Electronic document delivery 50
Document conversion 51
Scanning 51
Optical Character Recognition (OCR) 52
Document representation 52
Standard Generalized Markup Language (SGML) 53
Standard Page Description Language (SPDL) 53
Office Document Architecture (ODA) 53
Document transmission 54
File transfer 54
Message handling systems and document delivery 54
Employment 54

3.3 Applications for technical services 55
3.3.1 Shared cataloguing 57
LSP application protocols 58
Information Retrieval 58
Record Transfer 58
LSP implementations 58
Employment 59
3.3.2 Acquisitions 59
Electronic Data Interchange (EDI) 60
EDI and X.400 61
EDI and Libraries 61

3.4 Other library uses of research networks 62
3.4.1 Bulletin boards 62
3.4.2 Electronic conferences 62

4. Libraries and Research Networks Projects 65

4.1 Library and research network projects in North America 66
4.1.1 Library services on the Internet 66
Access 67
Directories 67
Connection 67
Login procedures 68

Terminal emulations 68
Search systems and interfaces 68
Information resources 69
Online catalogues 69
Non-catalogue information resources 69
Availability 70
4.1.1.1 Bibliographic utilities on the Internet 70
OCLC 70
OCLC's EPIC service 70
EPIC configuration 71
Technical services 71
RLIN 71
Internet access 72
RLIN databases 72
Technical services 72

4.2 Library and research network projects in Europe 73
4.2.1 The United Kingdom 73
4.2.1.1 Library services on JANET 73
Access 74
Directories 74
Connection 75
Login procedures 75
Search systems and interfaces 75
Information resources 76
Online catalogues 76
Non-catalogue information resources 76
Licensed databases 77
Gateways to online services 78
Availability 78
Interlibrary loan and document delivery 78
Technical services 79
Professional activity 79
Changes expected in 1991 80
4.2.2 Continental Europe 80
4.2.2.1 Sweden 80
Access 81
Information resources 81
Technical services 81
4.2.2.2 Norway 81
Access 82
Information resources 82
Technical services 84
4.2.2.3 Finland 84
LINNEA 84

LINNEA and FUNET 85
4.2.2.4 The Netherlands 85
The Pica database 85
The Pica network 85
Pica and SURFnet 85

4.3 Library and research network projects in Australia 86
4.3.1 Library systems on AARNet 86
Access 86
Information resources 86
Technical services 87

4.4 Library and research network projects in Japan 87
4.4.1 NACSIS and the Science Information Network 87
Access 88
Information resources 88
Technical services 88

5. Issues in Library Use of Research Networks 91

5.1 Technical issues 91
5.1.1 Library system/research network connectivity 91
5.1.1.1 Computer/Operating System Connectivity 92
5.1.1.2 Application Connectivity 93
5.1.1.3 Data structures 94
5.1.5 Transition to OSI 94
5.1.2.1 OSI over TCP/IP 96
5.1.2.2 OSI gateways for existing systems 97
5.1.3 Implementation of OSI Application Protocols 97
5.1.3.1 Z39.50 and SR 97
SR target systems 97
SR origin systems 98
5.1.3.2 FTAM 99
5.1.3.3 X.400 99
5.1.3.4 X.500 99
5.1.3.5 The ILL protocol 101
5.1.3.6 Virtual Terminal 102

5.2 Policy issues and other considerations 102
5.2.1 Governance 103
5.2.1.1 Governance in the U.S. 103
Research networks 103
Library networks/consortia 105

5.2.1.2 Governance in Europe and the U.K. 106
Research networks 106
Library networks 107
5.2.2 Intellectual property rights 108
Copyright 108
Copyright in a network environment 109
Network-accessible databases 109
Electronic document delivery 109
Imaging 110
Cataloguing 110
Authorship 110
International concerns 110
5.2.3 Appropriate use 111
Definition, application, and enforcement 111
Commercial usage and advanced services 112
Access to commercial information services 113
Access to bibliographic utilities 113
Electronic document delivery 114
Acquisitions 114
5.2.4 Convergence of library and computer centre operations 114
Areas of convergence 115
Administration 115
5.2.5 Impact on user policies 116
Management of system capacity 117
Authentication and access control 119
Methods of authentication and access control 119
End-user training 120
5.2.6 Impact upon staffing 122
Staff skills 122
General computer skills 122
Network-oriented skills 122
New positions 123
Increased demand for services 124
Reference service 124
Interlibrary Loan 125
End-user Training 125

5.3 Outlook for Developing Nations 125
Improvements through the use of research networks 126
Obstacles to research network use 127
Technological barriers 127
Economic barriers 127
Political barriers 128
Sociocultural barriers 128

6. Future Initiatives 129

6.1 The National Research and Education Network - NREN 129
6.2 The Coalition For Networked Information 131

7. Summary 133

Benefits of library use of research networks 134
Obstacles to library use of research networks 135
Role of the library

References 137
Relevant Standards 147

Appendix A: Global Research Network Topology A-1
Index to Research Networks A-67

Foreword

Research Networks and Libraries: Applications and Issues for a Global Information Network is the initial publication in the *UDT Series on Data Communication Technologies and Standards for Libraries*. The fundamental purpose of this series, derived from the mandate of the IFLA UDT Core Programme, is to promote the electronic transfer of data, both between libraries and between libraries and their users. The current focus of the UDT programme is the standards and related technologies, as well as the policy issues, which provide the basis for electronic resource sharing among libraries. At present, the results of five studies undertaken by the UDT Office at the National Library of Canada in 1990 and 1991 will be published in this series. In addition to the use of research networks by libraries, topics include electronic document delivery, organizational aspects of OSI for libraries, packet radio technology and EDIFACT standards. Additional publications will be added to the Series in accordance with future UDT investigations.

Research Networks and Libraries is a timely and informative report. The rapid emergence of research networks as one of, if not *the*, predominant themes for libraries in this decade, speaks to the need for this publication. This is a topic that presents innumunerable challenges to the library profession and may impact, in many ways, the basic nature of library and information services. This report views networks as a resource that can support the whole spectrum of library functions—from cataloguing and acquisitions to information retrieval and interlibrary loan.

This report is intended to complement the contributions of a variety of experts who have written on this subject. In reviewing the literature on research networks, however, one is struck by the overall lack of emphasis on a component that is critical to the successful technological exploitation of these networks by libraries—standards. *Research Networks and Libraries* emphasizes the role of standards, particularly those for Open Systems Interconnection (OSI), and attempts to illustrate the vast potential of research networks as a ubiquitous, global infrastructure for electronic data transfer.

To realize the full potential of research networks, libraries must bear the primary responsibility for planning information services in this new environment. Transforming the potential offered by research networks into actual library services will require the participation and direction of the library community.

It is hoped that this report will draw attention to the potential of research networks and, ultimately, mobilize the library community to adopt these networks as the basis for the evolution of electronic resource sharing.

Leigh Swain
Programme Director

Acknowledgements

Funding for this study was provided by the International Federation of Library Associations and Institutions (IFLA) as part of the 1990 Work Plan of the Universal Dataflow and Telecommunications (UDT) Core Programme. The support of the IFLA Programme Management Committee as well as the assistance of the UDT Advisory Committee and its Chairperson, S. Michael Malinconico, is also gratefully acknowledged. The author would like to express his appreciation to Leigh Swain, UDT Programme Director, for his overall direction of the study and to Paula Tallim, UDT Programme Officer, for her editorial direction and support.

The author would also like to acknowledge the written contribution of Joe Zeeman and Dennis MacKinnon, both of Software Kinetics, Stittsville, Ontario, in Chapter 2 and Section 5.1 of the study. Special thanks to Clifford Lynch, Director of Library Automation, University of California, Peter Stone, Information Services, University of Sussex, and Lorcan Dempsey of the Centre for Bibliographic Management, University of Bath, for reading an earlier draft of the paper and for many helpful comments and additional information. Thanks also to many individuals who answered numerous questions and supplied information that proved difficult to find elsewhere, including R. Gunderson, BIBSYS Project Manger, Norway, J. Huston, AARNet Network Technical Manager, Australia, C. Olausson, Librarian and Coordinator of Academic Library Development, Stockholm, Sweden, and E. Chapman of the RLIN Information Center, U.S.A.

Gary Cleveland
National Library of Canada

1. INTRODUCTION

A revolution is currently taking place in the world of computers and information technology that has the potential to transform the nature of communication and information exchange like the changes sparked millennia ago when primitive, unicellular organisms combined to form multicellular creatures—the interconnection, specialization, and communication among single cells allowed more complex life-forms with vastly greater capabilities to evolve. Similarly, the interconnection of widespread, isolated computers into large networks promises the emergence of a complex information infrastructure that will provide greater capabilities for the management and flow of information. Such enhanced capabilities will strongly impact many sectors of society that rely upon information, including research, education, business, industry, and government.

The basis of this information infrastructure will be *academic and research networks*, computer networks that have proliferated during the last decade linking the computing centres of universities, colleges, and research institutions throughout the world. They began in the late Sixties both as experiments in connectivity and to promote computer resource sharing among researchers (Larsen, 1990). Subsequent development of the networks focused largely on creating an underlying communications resource that allowed a variety of tasks to be performed between remote systems and users. These services include exchanging electronic mail messages, transferring files, participating in electronic conferences, accessing databases, and logging in to distant computers. They have become such a useful tool that they have spread throughout the research community connecting hundreds of institutions. The total number of individual research networks currently numbers in the hundreds worldwide, spanning North America, the United Kingdom, Continental Europe, Australasia, Asia, and the Far East (Quarterman, 1990).

Despite the ubiquity of research networks, and the usefulness of their services, the current level of functionality fails to exploit the full potential of these networks. As suggested by terms filling the literature on this topic, such as *virtual library* or *information superhighway*, the real value of research networks will be the vast amounts of information that will be available to end-users (Cline, 1990). By interconnecting decentralized information centres containing a variety of forms of information, end-users will be able to search and retrieve information relevant to their work directly from the computers in their workplaces. These computers will, in effect, become doorways to a worldwide library of electronic information.

Research networks and libraries

Libraries, whose central role is to organize and mediate access to information, are in a position to both contribute to, and benefit from, the development of an information network. Over the last thirty years, libraries have become dependent upon computer and telecommunications technologies to carry out many of their functions. These functions include not only those

involved in providing services to end-users, such as information retrieval, interlibrary loan, and document delivery, but also those that support end-user services, such as cataloguing and acquisitions (Arms, 1990c; Cline, 1990). Examples of telecommunications technologies and systems that libraries use in these areas range from terminal-to-host public catalogues and public data networks, to private electronic mail systems, vendors' electronic ordering systems, bibliographic utilities, and library networks. While libraries have realized significant improvements in their functions over manual methods, they have been left with a disjunctive set of telecommunications strategies to do their work; systems that accomplish essentially the same task often cannot communicate with each other, or have processes based on conflicting formats, thus limiting their functionality.

Research networks, because they are general communications resources interconnecting computers and users, have the potential to provide the infrastructure necessary to support a library's telecommunications needs. By employing these networks, and a number of application protocols and international standards discussed later in the paper, libraries can create end-user applications that can enhance both library services to end-users and the efficiency of internal library processes; research networks can provide greater access to the increasing amount of electronic information maintained by libraries, bringing a larger and more diverse set of information resources to a larger community of users, as well as bring the varied and often conflicting telecommunications strategies used to support all library processes into a more coherent framework.

Goals of the paper

The focus of this study is to provide interested librarians with a basic knowledge of research networks, focusing specifically on the manner in which they can be used by the international library community to support library functions and improve access to bibliographic information, as well as the technical and policy issues involved in research network use. While major projects involving library/research network connectivity are discussed, the primary focus of this study is on the potential for research network use and on available standards, rather on what is currently happening in the area. The strong momentum within the international academic community towards the development of this networking infrastructure, and the tremendous opportunities it offers to the library community, makes this topic of special importance to all librarians.

Research networks versus library networks

At the outset, it is important to point out that research networks are distinct from what is commonly known as networks among librarians. In the library world "network" has meant anything from the organizational resource sharing arrangements established among nearby libraries, to automated networks such as those operated by bibliographic utilities or the regional library networks that broker their services (Johnson, 1987). For the purpose of this paper, *library network* will refer to the latter category of automated network. While research networks developed within the academic community to provide the infrastructure for a variety of networking tasks, library networks developed specifically to support the nuts and bolts operations of libraries.

A point to point comparison of some of the differences found in the literature may make the distinction more clear (Arms, 1990a; Avram, 1988b; Lynch, 1988). In general, research networks can be characterized in the following way:

- they are general purpose communications infrastructures much like a highway system is a general purpose transportation infrastructure
- they were developed to provide the basis for a variety of networking tasks or end-user applications
- they allow communication among any number of computers at variety of research and academic institutions
- they operate in a peer-to-peer fashion, with each computer acting as an equal, or "peer" to every other (as distinct from terminal to-host networks where terminals act as simple input devices to centralized hosts)
- the network itself is independent of particular services offered
- they provide access to local and remote computing and information resources
- they are widely interlinked with other computer networks
- they support many services such as electronic mail, remote login, and file transfer
- they are exemplified by networks such as the *Internet*, JANET (Joint Academic Network), IXI, AARNet (Australian Academic Research Network), and NACSIS (Science Information Network).

Library networks, on the other hand, can be characterized in this way:

- they were developed specifically to support library technical processes such as shared cataloguing
- they provide a particular set of services and are operated by single organizations
- they are largely terminal-to-host networks
- the networks and the services they provide are indistinguishable
- they are not, for the most part, connected to other library networks, and thus are unable to exchange data (the exceptions are the few networks that are linked through the Linked Systems Project)

- each has its own custom system architecture
- they are exemplified by the bibliographic networks such as OCLC, RLIN, and WLN, as well as regional networks like the Illinois Library Network (ILLINET).

Organization of the paper

The paper is divided into six chapters, outlined below.

1. ***Introduction***

2. ***Overview of Network Concepts*** There is a dizzying array of terms and concepts which confront librarians when first exploring the world of research networks. This chapter introduces these terms and concepts including a discussion of network types, general network protocols, and specific application protocols of interest to librarians.

3. ***Library Applications of Research networks*** The areas within library science that have greatest potential for research network use fall into the domains of bibliographic information provision, document supply, and technical services. This chapter discusses applications within these areas, discussing the impetus for each area's use of research networks, and the application protocols and standards that are most appropriate for supporting the applications.

4. ***Library and Research Networks: Projects*** There are currently major projects linking libraries and research networks underway in the United States, the United Kingdom, the Netherlands, Sweden, Finland, Australia, and Japan. This chapter outlines these projects and discusses a number of patterns that appear in their use of research networks.

5. ***Issues in Library Use of Research Networks*** A number of major issues must be resolved before library applications of research networks can be fully developed. These include both technical issues and softer issues such as governance and copyright. This chapter discusses some of the major technical and political problems that will have to be overcome in the process of creating information networks.

6. ***Future Initiatives*** The evolution of research networks is progressing throughout the world very rapidly. Despite major advances, the development of these networks is still at a very early stage. This chapter reviews some future directions and major initiatives that are underway to create an information infrastructure.

7. ***Summary***

Appendix A. ***Global Research Network Topology*** Research networks are actually a patchwork of networks of differing characteristics rather than a homogeneous system. This chapter provides an overview of current global research network topology, outlining the international availability of these networks and their links to other networks. It is intended to make librarians aware of the great extent to which research networks are currently used worldwide, rather than serve as an up-to-date directory of research networks.

2. OVERVIEW OF NETWORK CONCEPTS

What is meant by the term "network"? The *Oxford Dictionary of Current English* defines a network as an "arrangement with intersecting lines and interstices, complex system of railways, etc., chain of interconnected persons or operations or electrical conductors, group of broadcasting stations connected for simultaneous broadcast of same programme". While this definition is somewhat obsolete in that it does not refer to networks used by or with computer systems, it does capture the essence of any kind of network; that is, a set of objects that have complex interconnections allowing some common activity to be carried out. As the definition suggests, original networks were railway networks, and a surprisingly large part of networking terminology descends from railways, generally via telephony and broadcasting, such as *switching*, *station*, *terminal*, *routing* and *line*.

Data communications via networks were originally developed to maximize the use of expensive computer resources, such as large mainframe computers or large amounts of computer storage. Research networks were created expressly to provide researchers with access to one or more supercomputers and resource sharing continues to be an important use of the networks. A number of the network services described later in this chapter enable resource sharing, either by letting a remote computer appear to be a local terminal, or by designing special network applications which allow two or more computers to cooperate in accomplishing certain tasks.

Some definitions

It is important to initially make some distinctions about what is exactly meant by the term *computer network*. This term has been used to refer rather broadly to:

- a set of terminals linked to a central computer, known as a *terminal-to-host* network
- a set of computers linked with one another, known as a *peer-to-peer* network
- the actual physical network underlying both terminal-to-host and peer-to-peer networks that provides the pathways for communication, known as a *data network*.

These three terms will be described in more detail below, beginning with the physical network, or data network, that forms the basis of all types of networks.

Data networks The *data network* is the actual physical system that underlies all network communications. It is the set of physical connections that support the movement of the 0's and 1's, or *data*, between end points. The media supporting the connections varies from network to network ranging from standard telephone-style copper wire, to coaxial cable and fibre-optic

cable, through to radio and microwave signals. The way data is transmitted over the data network can also vary, with some networks dividing data into small parcels, called *packets*, and others sending a continuous stream via a *circuit*. What is actually found at the end points of the underlying structure does not matter. The end points might be a terminal connected to a computer as in a terminal-to-host configuration, or a computer connected to another computer as in a peer-to-peer configuration.

While most of the physical differences between data networks are both invisible and irrelevant to end-users of a network system, they do have some impact. Specifically, differences in underlying technology can affect the range and quality of services provided, the ability of one network to communicate with other networks, and the overall speed of data transmission.

Terminal-to-host networks This type of network uses the data network to connect "dumb" terminals with little or no processing power to a central computer host. The earliest terminal-to-host networks were created to allow the resources of a single machine to be shared. This was a necessity when computers were scarce, very large, and very expensive. All that was required was a simple device at the user end to allow the input of instructions to the computer, a physical connection to transmit the instructions, and often a device to allow the computer to generate output at the user's location. This interaction is based on the *master-slave* paradigm; the user interface that the user sees, the software that manipulates the data, and the data itself, are all located in the central host, the "master" of the system. The local terminals are "slaves" to this host, and act simply as input/out devices for the activity of the host. Where actual computers are used to connect to these central hosts, the local computer to pretends to be, or *emulates*, a simple terminal. In this configuration, the processing power of the local computer is not used.

The status of PCs in terminal-to-host networks was eventually raised from slaves to servants. In *master-servant* relationships, some of the tasks originally located in the remote host were carried out by the PC. Terminal-to-host networks are typical of most library networks familiar to librarians such as RLIN and OCLC in the United States, BLAISE in the United Kingdom and UTLAS in Canada, as well as commercial information services such as DIALOG or BRS.

Peer-to-peer networks A *peer-to-peer* network is a network of independent computers that are capable of communicating and sharing resources, with the full processing power of each computer available for use. Each computer is considered to be equal to any other computer regardless of type or size. While a peer-to-peer network requires an underlying data network, it is intended to accomplish more than just the transfer of data. They are general communications infrastructures that manage and support any number of tasks between remote systems and users such as electronic mail, remote login, file transfer, remote job entry, and conferencing.

The procedure by which two computers cooperatively accomplish a task is based on a *client-server* paradigm. The computer at the user's end of the network acts as a *client* of the services provided by the remote computer. The client system sends messages containing requests which are responded to by the server system. These requests may be for any kind of computing service, such as a file transfer, a database look-up, or an online transaction. The advantage of

this approach is that it separates the communication aspects of cooperation from the machine processing aspects, so that the client computer system need not be aware of the details of how the server system stores or processes its data. Mail servers, file servers, and database servers, are among the types of servers that may be available on a peer-to-peer network.

Peer-to-peer networks can encompass networks from two or three microcomputers on a local area network (see 2.1.1), to the largest of research and commercial networks, such as Digital Equipment Corporation's internal EASYnet which connects more than 34,000 computers around the world (Quarterman, 1990).

While peer-to-peer networks have been created for a variety of purposes, such as the support of military activities and private business, research and academic networks are the type of peer-to-peer network that are the focus of this paper. As outlined in the introduction, research networks link together hosts residing at research institutions with the purpose of facilitating collaboration, resource sharing, and information exchange among researchers, educators and academics. See Appendix A for descriptions of major research networks worldwide.

The remainder of this chapter introduces some of the terms and concepts with which a librarian might find him or herself confronted when exploring networking issues. No attempt is made in any of these descriptions to give a detailed technical account of the technology or standard; the intention is rather to give the reader an introduction to both the concepts and the terminology used in the world of research networks.

2.1 Network types

One broad means of classifying networks is by the geographic area which they are intended to cover; a network which is intended only to link computers in one room or in one building has very different requirements from a network intended to link computers in different continents. This classification is one of Local Area Network, Metropolitan Area Network and Wide Area Network. Each of these is discussed further in the following sections.

2.1.1 Local Area Networks (LANs)

A LAN is a network that spans a small area, typically a single building or site. Most LANs allow centralized use of network resources, such as data and software. There is often a central *server* on which all or most data or software is stored. When a user at an individual computer on the network wishes to create, examine, or update data, it is retrieved by the local computer from the server, processed by local software, such as a word processor or spreadsheet (often initially loaded to the local computer from the server), and then saved back to the server.

Such an arrangement offers a number of benefits. It maximizes the use of the processing power of the desktop computer, while minimizing the need for expensive resources, such as large

amounts of data storage or sophisticated printers, to be provided for each user. It allows formal, organizational procedures to be established for ensuring the integrity and security of data by storing it centrally. It also allows many users to have access to the same software and data, although not necessarily simultaneously.

What sets a LAN apart from other networks is that the technology used for physically transferring data in most LANs imposes maximum limits on either the physical size of the network and/or the number of systems which can be connected to it. These networks are typified by high data transfer rates, (generally from 2 to over 100 Mbps million bits per second), very reliable data transfer, and connectionless communication protocols (see 2.2.2).

LANs also normally use *broadcast* or *multicast* technology: every computer on the LAN receives all the data transferred over the LAN, regardless of where it is intended to go. Only the computer to which the data is addressed acts on it. The other computers ignore data not addressed to them, or simply pass it on.

The major topologies for LANS are a *ring topology*, a *star topology*, and a *bus topology* (see Figure 1). In a ring topology, every network station is connected to two neighbours and the network has no logical or physical end. Data moves sequentially, from one station to the next. In a star topology, a single central hub station is connected to each node. The computers at the points of the star then communicate through physical or logical connections in the hub. In a bus topology, data is broadcast; that is, every station on the network receives all data more or less simultaneously. Only the station to which the data is addressed acts on it, other stations simply ignore the data.

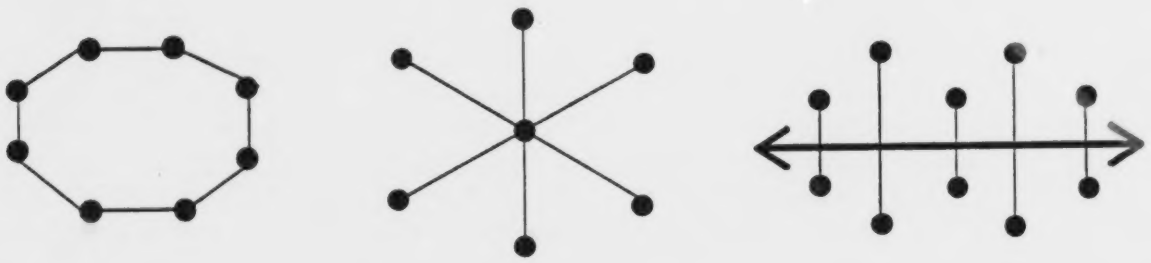

Figure 1- Network Topologies

These topologies give rise to the physical constraints on LAN size. In a ring topology, the data is returned to the original sender within a certain length of time to allow the sender to be certain that the data was correctly delivered. Since the electrical signals in the cable travel at a finite speed, the length of time allowed for the return of the message determines the maximum size of the

network. With a bus topology the requirement for the data to be received more or less simultaneously by all stations determines the maximum size of the network.

Many LANs employ one of two broad strategies to ensure reliable data transfer, *token passing* and Carrier Sense Multiple Access/Collision Detection, or CSMA/CD. In token-passing networks, an electronic token is passed as a marker sequentially between stations, with only the station in possession of the token permitted to transmit to the network. In CSMA/CD, any station on the network is permitted to transmit whenever the communications channel is idle, but stations are able to detect when two or more stations transmit simultaneously (leading to corrupted data) and follow prescribed rules for retransmitting.

Ethernet (IEEE 802.3) is the most widely available and best known CSMA/CD network. It uses a bus topology. Token-based networks include IBM's Token Ring network and the Manufacturing Automation Protocol (MAP) which uses a token-passing bus architecture as specified in ISO 8802-4.

2.1.2 Metropolitan Area Networks (MANs)

Metropolitan area networks may be thought of as those networks that are too large to be treated as a single LAN, but too small to be a wide area network, although it is very difficult to draw a clear distinction between a MAN and WAN. Metropolitan networks are generally under the authority of a single administration, such as a university or a library system, with all the component systems under the authority of the same administration. They are generally within a single locality such as a city, offer higher data transfer speeds than WANs but lower than LANs, and use the same architecture and protocols as WANs. IEEE standard 802.6 addresses Metropolitan Area Networks.

2.1.3 Wide Area Networks (WANs)

Wide Area Networks, as the name suggests, span large geographic distances, from a few miles to the area of a continent. There are important differences between WANs and LANs. First, there are few physical constraints on the size of WANs. Secondly, WANs have until recently operated at significantly slower data transfer speeds, ranging from 300 bps to 1.54 Mbps. Presently, some of the major network links are being upgraded to offer speeds of up to 48 Mbps, and there is discussion of even higher speeds in the future. A further distinction is that, while computers and terminals normally attach directly to the LAN communications medium such as a coaxial cable or copper wire, WANs normally consist of special purpose computers that are interconnected by dedicated communications links. Transferring data over a WAN generally involves the use of one or more of these special computers. Finally, WANs normally introduce some delay between transmission and receipt of data whereas LANs introduce very little delay (Comer, 1988).

Unlike most LANs, in which the data passes through every station in the network, data transferred across a WAN is sent more or less directly from the sending node to the receiving node, via one or more intermediaries, if necessary. Networking software for WANs must

therefore address a problem which does not arise in a LAN: how to determine the best way to move data from one node to another, a process known as *routing*. In large networks, where there are many possible paths between nodes, routing may involve complex calculations taking into consideration both time and cost. To accomplish routing most efficiently, routing software must be provided with up-to-date information about the state of network connections.

Whether they provide connections between one computer and another or between terminals and computers, networks supporting wide-area communications can be divided into two basic types: circuit-switched and packet-switched.

2.1.3.1 Circuit-switched networks

Circuit-switched networks operate by forming a continuous connection (a *circuit*) between the two ends that are in communication. In its simplest form, a circuit-switched network establishes a continuous, dedicated physical path for each circuit. The advantage of circuit switching lies in its guaranteed capacity; once a circuit is established, no other network activity will decrease the capacity of the circuit. A disadvantage, however, is that they may be relatively cost inefficient. Circuit costs are set according to elapsed time and are independent of the rate of data transfer; the same rate of charge applies whether the circuit is idle or not. When the full capacity of the circuit is constantly used, however, a circuit-switched network may be very cost effective.

The telephone system is by and large a circuit-switched network. For this reason, circuit-switched data networks often make use of the telephone system, using special equipment (modems) at either end of a telephone line to convert the data from digital, the form in which it is used by the computer system, to analog, the form carried over the telephone system.

Leased Lines

Leased lines are permanently open switched circuits that have been rented from common carriers. Such circuits are available with the widest possible data transfer rates, ranging from 1200 bps to over 44 Mbps and higher. Leased lines, however, are inherently inflexible as they provide only point-to-point connections. If a different connection is required, then a different leased line is needed. Dedicating the whole capacity of a physical circuit to the use of a single pair of communicating computers is, however, inefficient because the circuit is idle most of the time. Two computers linked through leased lines do not constantly transfer data. Instead, one system is generally waiting for the other to do something, such as for an operator to press a key or for the computer to process an instruction or retrieve data. Leased lines, however, have been widely adopted in the United States because they have been advantageously priced.

2.1.3.2 Packet-switched networks

The physical circuit could be used much more efficiently if the data of more than one pair of communicating computers could be carried over it. This is what packet switching accomplishes. In a packet-switched network, data is divided into small segments, called *packets*, which are transferred using high-capacity inter-machine connections. For example, a file to be transmitted

between two machines may be broken into many packets that are sent across the network one at a time. Each of these packets is put into an "envelope" which has all the information needed to get it to the correct destination and to allow the original file to be put together again. These packets may be intermingled with packets from other sources heading to other destinations on the same physical cables of the network. The network delivers the packets to the specified destination where network software reassembles them into a single message again.

There are two kinds of packet switching, *virtual circuit* (or connection oriented) and *datagram* (or connectionless). In a virtual circuit, the first packet is used to establish a connection and all subsequent packets seem to follow the same route; the network service guarantees that all packets will be delivered to the recipient, in the correct order.

In a datagram network, each packet is routed independently, and different packets may follow different routes; the network service does not guarantee that all packets will be delivered in the right order, nor even that all packets will be delivered. Because some routes will take longer than others, it is possible that later packets may arrive before earlier ones. It is the responsibility of the users' software to ensure that packets are reassembled in the correct order.

X.25

Rules governing the definition and use of packets are contained in CCITT recommendation X.25, the international standard for packet switching. X.25 uses packet switching to provide virtual circuits between nodes in a network. A companion standard, X.75, provides virtual circuits between dissimilar networks. The standard specifies both how packets are composed, and the procedures for sending and receiving them. The composition of a packet includes such things as the permissible length of data, the structure of the "header" prepended to the data for routing and error detection purposes, and the algorithm used to generate check values. The procedures address such matters as placing and terminating a call, controlling the flow of packets, and error handling.

X.25 networking is increasingly being used to replace circuit-switched networks. This is often accomplished by means of a Packet Assembler/Disassembler (PAD), which allows standard terminals to connect to their host computers over X.25 networks. RLIN, for instance, has replaced its large circuit-switched data network with an X.25 network in which the terminals communicate with a local or remote PAD which then communicates directly with the RLIN network processors (Richards & Lerche, 1989).

2.1.4 Public Data Networks (PDNs)

A public data network is a network that is publicly accessible for a fee and that provides basic services upon which users can build their own networking applications. Outside North America, such networks are usually run by national telecommunications agencies, often the same bodies as are responsible for postal and telephone services (PTTs).

Since the approval of the X.25 standard in 1976, increasing numbers of PDNs based on packet-switched networks have become available. The first such network was the Canadian *Datapac* network, established in 1976. American PDNs include *Telenet*, which claims to be the world's largest, and *Tymnet*, which claims to be the most ubiquitous, with 750 US locations and connections to 68 countries. In Europe, PDNs are commonly known as public packet-switched data networks (PSDNs). European PDNs include British Telecom's PSS and IPSS services, France's *Transpac*, and Denmark's *Paxnet* and *Datapak*.

2.1.5 Value-added networks

Many PDNs are evolving into networks which provide services which build on the basic data transfer capabilities of the network to give subscribers additional benefits from the network for an additional charge, including access to commercial information utilities, electronic mail, and airline timetables. The Canadian PDN, *Datapac*, for instance, offers a number of added value services, including the Envoy electronic mail service and the iNet 2000 database access services. In the United Kingdom, British Telecom's Telecom Gold provides similar value added services, as do many of the PDNs in the United States, such as *Telenet*, *Dialnet* and *Tymnet*, all of which provide, among other services, access to the DIALOG databases.

2.1.6 Integrated Services Digital Network (ISDN)

The Integrated Services Digital Network (ISDN) is a set of communications standards aimed at the creation of a global telecommunications network supporting a wide range of voice and non-voice applications. For the user it offers the possibility of carrying both voice and data simultaneously over a single physical connection, providing an inexpensive, universally available data network.

There are two kinds of physical transmission systems used in the telecommunications infrastructure: 1) channels designed to carry an analogue signal corresponding to the human voice, and 2) channels designed to carry a digital signal directly. Both voice and computer data can be carried on either analogue or digital channels, however, carrying data over analogue channels requires additional pieces of equipment, called modems, at either end. Using analogue channels for data is therefore more expensive and inherently less reliable than using digital channels because two additional data conversion processes are required.

An increasing proportion of the telecommunications infrastructure worldwide, particularly in North America and Europe, is being changed from analogue to digital facilities, both for transmission across cables and for switching at exchanges. Yet, the last leg of most circuits, typically to the subscriber's telephone or modem, remains an analogue channel. One of the aims of ISDN is to extend the digital network to the end-user.

While the standards involved in ISDN are both numerous (78 CCITT Recommendations) and complex, two fundamental levels of service have been defined. The *basic access* service provides a user with two 64 Kbs B channels (bearer channels) for voice or data, plus one 16 Kbs

D channel (data channel) for control and signalling. *Primary access* service provides a user with 23 64 Kbs B channels and one 64 Kbs D channel (for a total of 1.536 Mbs).

The following principles underlie the development of ISDN (Martin, 1988):

- ISDN must evolve from the existing telephone networks of the world
- new services that are to be offered must be compatible with the basic 64 kbs switched digital channel
- a worldwide ISDN will probably not be fully operational until the end of the 1990's, although some ISDN services are already available in some areas
- during the transition period, ISDN will rely heavily on interconnections between national ISDNs and existing public data networks
- ISDN will implement intelligence in order to perform such functions as maintenance, system control, and network management
- ISDN will use a set of layered protocols that conform to the OSI Reference Model.

It must be stressed that ISDN involves no new technology; all the technology required to implement ISDN already exists and is in use in telecommunications systems. ISDN is rather an exercise in standards creation and implementation, although it is severely hampered by the fact that North America has standards for digital telephony that are incompatible with those used by the rest of the world. Furthermore, there is some doubt as to whether basic ISDN service will provide sufficiently high-speed communications for the data-intensive applications which are currently in development or are expected. For example, the transfer of graphic images may be expected to require data transfer rates significantly greater than those offered by the ISDN basic access service. To help overcome this difficulty, additional data services are being standardized, including Wideband ISDN (W-ISDN) and Broadband ISDN (B-ISDeN).

Nevertheless, the introduction of ISDN may be expected to have a significant impact on data communications. For one thing, communication using ISDN, which charges on the basis of connection time, is likely to be considerably cheaper than communication using a packet-switched PDN, which charges on the basis of volume of data transmitted (Tuck, 1989).

2.1.7 Store-and-Forward Networks

Store-and-forward networks, unlike circuit-switched or packet-switched networks, do not maintain a constant connection between computers while accomplishing tasks. Most often used for interpersonal communications and some kinds of resource sharing (e.g. remote job entry), they do not require real-time interaction between the local system and the remote system. With electronic mail, for example, the sender can transmit a message without the need to wait for a

reply, while the receiver does not have to be present to receive the message. This is analogous to conventional mail in which the sender posts a letter with the expectation that it will eventually be delivered and that whatever action is required by the letter will be taken.

Advantages of store-and-forward networks are that communications costs are minimized, for example, by storing all out-going mail until cheap rates became available. Equally, if there are multiple stages involved in the delivery of mail, each intermediate system could also store the message until there was a convenient time to forward it to the next system in the delivery chain. Disadvantages are that services are limited to tasks that are inherently non-interactive, such as email. Services like interactive resource sharing, remote login, or directory look-up, are not possible. An example of a common store-and-forward network is BITNET, a network dedicated primarily to email.

2.1.8 Internetworks

An internetwork, or internet, is a network of networks. Individual networks generally use a single technology, such as Ethernet or X.25, to link computers and to allow the direct mutual exchange of data. An internetwork, in contrast, is composed of multiple networks, each of which may use different technologies. The individual networks are connected by computers that exist as a *node* on at least one other network. These common nodes are known as network *routers*; data from a computer on one network passes to a computer on another network through a router. Large internetworks, such as the American *Internet*, are unified by the use of a single protocol suite. In the case of the *Internet*, the common protocol is TCP/IP (see 2.2.4.3). All the applications supported by this protocol suite will normally be available to any computer on the *Internet*. A European example of a internetwork is IXI, an OSI-based network that interconnects many research networks throughout the Continent.

It is also possible to interconnect networks which use incompatible protocol suites. This is done by means of *gateways* which translate data from one protocol to the other. Incompatible internetworks may be connected in this way, and the result is often called a *metanetwork*. The applications available across a metanetwork are generally very limited, often restricted to electronic mail.

2.2 Overview of protocols

In the diplomatic world, a protocol is a set of formal rules specifying how business is to be conducted between sovereign states. These rules are designed to ensure that communication is not impeded by misinterpretation of words or deeds because of cultural or linguistic differences. The rules of protocol cover the whole of diplomatic life, from the generalities of maintaining relations to the minutiae of who sits where at a formal dinner.

A communications protocol serves precisely the same function. It is a set of formal rules that specify how to conduct business between independent computer systems. These rules are equally designed to ensure that communication is not impeded by misinterpretation of words or deeds because of system differences. The communications protocol will specify both the broad terms of what services one computer system can expect to be available from another and also the minutiae of what kinds of data may be exchanged together with its order and syntax.

2.2.1 Layering, transparency, and virtual facilities

There are three concepts used in discussing communication protocols and architectures. These are *virtual*, *transparency*, and *layering*. The first two concepts can be summarized as follows:

- something that is virtual appears to exist, but in fact does not
- something that is transparent actually exists, but does not appear to exist. In other words, it is invisible.

While this may be a little vague at first, a concrete example illustrating these two concepts may be helpful. When two people have a conversation, they are generally only aware of the flow of thoughts moving back and forth between them, and not of the underlying mechanics that allow the flow to take place, such as the moving of lips, sound waves striking ear drums, or the intricacies of grammar. To the participants, the flow of thoughts is like a meeting of minds, though their minds do not actually meet. The conversation, therefore, can be said to be a *virtual* meeting of minds. Conversely, awareness of the actual physical and grammatical mechanics of the conversation is not necessary for it to be successful; the actual mechanics of the intercourse can be said to be *transparent* to the conversants. The use of these concepts with regard to communications protocols will become clear in the discussion of layers. Suffice it to say that they describe opposite approaches to understanding what is accomplished through layering.

Layering

Layering is the division of a total communication activity into a hierarchical set of smaller activities, allowing each layer to be treated separately. In the example of the conversation, grammar can be treated as a distinct layer as it is a clearly distinguishable part of the process of human communication with its own definable rules. Other layers can include the speech system, defined by the principles of physiology, as well as layers of biology and of physics, also defined by their own sets of rules.

The layers in communication protocols, while there are several, are divided broadly into two classes:

- *network* or *lower layer* protocols, or those that deal with the problems of getting the data from the originating system to a destination system

- *application* or *upper layer* protocols, or those that deal with the problems of allowing one computer system to effectively interoperate with another.

While network protocols will generally form part of the operating system of a computer system, or may even be embedded in the hardware, application protocols will generally form part of the application software (the program) which is available to the user of a computer system.

Layer transparency An important aspect of layering is that the activities that take place at lower layers are transparent to the layers above. In the conversation example, the event taken at the level of thought does not need to take into account the workings of grammar. Similarly, the level of grammar has its own rules and does not need to take into account the workings of the speech system. The speech system, in turn, has *its* own rules, and does not need to take into account the workings of biology. This same transparency of lower layers continues down through to the lowest layer that is under consideration. Although the activities at all layers are present in the conversation, only the highest level is of concern to the participants.

In computer terms this means that an application, such as file transfer, does not need to specify how the data physically gets transmitted over a wire, nor how the connection between two systems is established and maintained. It only needs to pass on to the layer below the data destined for the remote system, as well as the instructions that the lower layer needs in order to be able to accomplish its task. Equally, the lower layers do not need to know anything about the application. They simply receive data from the layer above, add the information required for their layer, and pass it on.

Layer envelopes Implementation of the layered architecture is achieved by enclosing information in "envelopes". The software at each layer receives data in an envelope from the layer above, and without altering it in any way, takes actions needed at the current layer and adds any necessary data of its own. It encloses the entire package into another envelope, and passes it to the layer below. The overall effect is analogous to that of a set of Russian dolls. Just as each doll is enclosed completely within a larger doll, so are protocol layers contained within another. In the transmitting system, the data of each protocol layer is enclosed within the data of the next lower layer. At the receiving end, the envelopes are then removed layer by layer until the software at the highest layer finally receives the message that was originally sent. The higher layers never see the envelopes which the lower layers put around their data.

Layer independence Another important principle in layering is layer independence, the ability to modify the activities at any one layer without affecting the specification of the layers above and below. In this way, the same set of higher layer protocols can be implemented over various lower layer architectures, such as a LAN or an X.25 network, without the necessity to make changes in the higher layer protocols.

Advantages and disadvantages of layering The use of a layered architecture offers a number of advantages:

- it produces shorter and more intelligible protocol specifications because each one deals with only one aspect of the general problem
- it allows the appropriate expertise to be assigned to each aspect of the problem
- it allows a specific communications problem to be solved and then applied to different circumstances and environments
- any given layer can be modified or upgraded without affecting the other layers.

Layered architecture has a number of costs, as well. Specifically, layered architecture results in:

- ***increased complexity of processing***. Each protocol layer may require additional processing over and above what is strictly necessary for the immediate situation. This can result in disappointing response times and/or in a need for more powerful computers
- ***higher telecommunication costs***. Each protocol layer adds data to the message to be transmitted, thus increasing the size or number of data packets, and therefore the cost. For an application which sends many small messages, the additional cost imposed by layered communication protocols may be significant
- ***the potential need to duplicate some functions in different layers in order to guarantee layer independence***. This may add further to processing complexity.

2.2.2 Modes of Communication

One of the most controversial areas in the design and specification of network protocols has been that surrounding *connection-oriented* and *connectionless* services. While it is obvious that there must be a connection between two computers at some point in order for a transfer of data to take place, there is considerable scope for disagreement as to which layers can be assumed to establish and maintain such connections.

The argument centres around the extent to which applications on a network can assume the existence of a reliable connection. A reliable connection is the ability of one party to set up a connection, send data with the assurance that it is transferred accurately in the proper sequence, and to conclude a connection. If an application can assume the existence of a reliable connection, then its design may be quite different from that of an application which cannot assume a reliable connection. In this case, the application must assume that every message sent needs to be placed in a correct context and sequence by the receiving application.

2.2.3 The OSI Reference Model

A somewhat elaborate description of the OSI Reference Model is given here because it is representative of layered protocols in general and is the most detailed of them. In addition, OSI applications are expected to become widespread in library environments over the next several years.

The OSI protocols are developed by either the International Organization for Standardization (ISO) and the International Electrotechnical Committee (IEC) or the *Comité consultatif international de télégraphie et téléphonie* (CCITT) of the International Telecommunications Union (ITU). They are issued as ISO/IEC Standards and as CCITT Recommendations, the texts of which are virtually identical.

OSI uses a 7-layer architecture, known as the OSI Reference Model. Each layer is briefly described in below.

- ***Physical layer*** The lowest layer is the Physical layer. It is at this layer that bits of data are passed from a computer to the wire or other physical connection between machines, and then back from the wire to a computer. This layer does not include the physical medium itself. Physical transmission media are not part of OSI.

  The Physical layer forms the interface between a particular physical medium and the computer. The details of conversion of digital data into the analogue form often used by communications media are hidden from the upper layers as are details of physical connectors, pin assignments, voltages, waveforms, digital-to-analog converters, and so forth. There are many OSI protocols at this layer, corresponding to the many different kinds of transmission media.

- ***Data link layer*** The second layer, the Data Link layer, accomplishes the error-free exchange of data between two directly connected systems. It is thus mainly concerned with the detection and correction of errors arising during transmission. A number of protocol standards relate to this layer, catering to the various real networking environments, including those designed for LANs, packet and circuit-switching networks, the telephone network, and ISDN.

- ***Network layer*** While the first two layers handle exchange of data between directly connected, adjacent systems, the third layer, the Network layer, provides services that enable systems anywhere to exchange data. It handles the routing of data from originating system to destination system via any intermediate systems and makes it unnecessary for the upper layers to know the actual paths taken between the systems. The services at the Network layer also includes the transfer of data between networks.

- ***Transport layer*** The Transport layer, layer 4, is the first in which data is transparently exchanged between end systems; all the lower layers may involve the processing of data by intermediate systems to ensure, among other things, proper routing and error correction.

Above the OSI transport layer are the three upper layers of the OSI model: the *Session layer*, the *Presentation Layer* and the *Application Layer*.

- ***Session layer*** The Session layer adds structure to the connection provided by the Transport layer. While the transport service provides a reliable connection for the transfer of data, the Session layer manages the communication session over that connection. For example, it can structure a session into a half- or full-duplex dialogue.

- ***Presentation layer*** The Presentation layer ensures that information is transferred in a form mutually intelligible by different, possibly incompatible computer systems. For instance, one computer system might store integers as a series of bits with the most significant bit first, while another system may store integers with the least significant bit first. The Presentation layer ensures that data will be transferred between these systems in a form both can understand.

- ***Application layer*** The Application layer is the highest, and most complex, layer of the OSI model. It is important to note that the Application layer does not specify what applications are to be available to a user of a computer system, nor how any particular implementation should appear to a user. Rather, the Application layer specifies what communications services are to be available to a computer system for a number of specific purposes. For instance, the *File Transfer, Access and Management* (FTAM) protocol specifies the communication services which support operations involving files. These services can be used by any computer application that has a requirement to read, write or otherwise manipulate files across a network.

  Application layer protocols can be used in combination to provide any given computer application with the communication services it requires. Thus, there is no need to specify an indefinite number of protocols for each specific application. Instead, the combination of a few more or less generic protocols will normally be sufficient. Some of these protocols may be available for direct use by an end-user program, whilst others, which provide common services, will only be used by other protocols and the end-user will be largely unaware of their existence.

2.2.4 Major protocol suites

A major impediment to global networking lies in the fact that multiple network protocol suites have been specified and implemented. These suites are, on the whole, incompatible with each other. This variety of protocol suites has come about from a number of causes:

- the need for vendors of computer systems to protect their market share leading to the development of proprietary protocols
- delays inherent in the international standardization process leading to the adoption of *ad hoc* interim protocols
- differing points of view on technical issues leading to the creation of incompatible protocols.

While the present situation is likely to change with the increasing maturity and stability of the OSI protocol suite, the transition to general use of OSI protocols will be gradual at best. Systems designed for interworking now must recognize the fact that multiple protocols will be in existence for at least the next decade.

The following sections briefly describe the most widely used of the various protocol suites. They are not described in detail, however, because such detailed descriptions are beyond the scope of this overview. Readers who require further details are advised to consult the references cited for each suite.

2.2.4.1 The OSI Protocol Suite

The OSI protocol suite is based on the OSI Reference Model as described above. There is not, however, a one-to-one correspondence between layer and protocol. There are at least two protocols at every layer (except the Physical), to support the connection-oriented and connectionless modes respectively; at the physical layer, where a separate protocol is required for each different communication technology; and the application layer, where a different protocol is required for each different aspect of communication (MacKinnon *et al*, 1990; Rose, 1990). The following briefly describes the general classes of protocols at each layer.

Physical and Data Link Protocols OSI supports a large number of protocols at the Physical and Data Link layers to allow the use of a number of networking technologies including the main LAN technologies, X.25 packet switching, switched circuits using modems, and ISDN.

Network Protocols The network layer is one of the most complex of OSI's layers. It is the responsibility of the network layer to convey data from end-system to end-system over possibly very complex network configurations, regardless of the underlying communications technology. Thus, communications must be supported over connectionless technologies such as LANs, as well as over connection-based technologies such as X.25 and ISDN. In addition, the network

layer must be able to relay data from one system to the next determining the most appropriate route for data to take, although this intermediate-to-intermediate system (IS-IS) routing is not fully specified.

The Network layer is therefore the only one of the lower layers for which an internal structure has been explicitly defined. The principal protocols at the network layer are the X.25 Packet Level Protocol, the Internetwork Protocol, and the ISDN Network Protocol.

Transport Protocols The principal protocols at the transport layer are the Connection-Oriented Transport Protocol, which can operate using either the connection-oriented or connectionless network service. There is also a Connectionless Transport Protocol, normally used with the connectionless network service.

The connection-oriented transport protocol defines five classes of service, corresponding to different levels of reliability of the underlying network connection. Transport protocol class 4 is broadly equivalent to the *Internet*'s TCP protocol.

Session Protocols The session layer is implemented by one of two protocols, the Connection-Oriented Session Protocol and the Connectionless Session Protocol. The services provided by the connection-oriented session protocol include full or half-duplex operation, orderly release and abortion of a connection, and the insertion of synchronization points.

Presentation Protocols There are two protocols in the OSI Presentation layer, the Connection-Oriented Presentation Protocol and the Connectionless Presentation Protocol. They support the identification of the syntaxes to be used for transferring data. The principle standard for representing data is the Basic Encoding Rules specified for use with data structures defined using the Abstract Symbolic Notation One (ASN.1).

Application Protocols There are two broad classes of application protocol: those which define services for specific applications, such as interlibrary loans or interpersonal electronic mail, and those which define services that are generally useful in a variety of applications, such as establishing and maintaining communication between remote systems, ensuring the integrity of database operations, and handling the management of distributed applications. The major application-specific protocols are discussed in later sections. A brief description of the common *services elements*, relatively small, simple protocols providing specific communication services, follows here.

The Association Control Service Element (ACSE) is responsible for creating, maintaining, and terminating an *association* between two independent computer systems to allow them to interoperate. All connection-oriented applications use the ACSE.

The Remote Operations Service Element (ROSE) allows one computer system to interactively execute procedures, or *operations* on another system. This protocol is an important one, as it is central to many application protocols, including the Directory and Remote Database Access.

The Commitment, Concurrency and Recovery Service Element (CCRSE) ensures the integrity of data in a distributed application, such as updating a distributed database, by requiring that each system can guarantee the success of an action, such as a file transfer or database update, before that action is actually carried out.

2.2.4.2 TCP/IP

This abbreviation for Transmission Control Protocol/Internet Protocol identifies the suite of protocols initially developed for the U.S. Defense Advanced Research Projects Agency from the mid 1970's (Comer, 1988). It is today the most widely used set of standardized, vendor-independent protocols. This widespread use stems partly from the fact that TCP/IP was incorporated into a version of the UNIX operating system and, by this means, became readily available to the academic and scientific communities who were the largest users of UNIX systems and most in need of networking services. Further, products implementing TCP/IP support communications among more hosts and different types of data links than any other protocol suite. The types of hosts connected through TCP/IP range from Cray supercomputers to IBM PCs, and Apple Macintoshes. TCP/IP can be carried over Ethernet, optical fibre links, DECNET, X.25 networks, satellite links, and hi-speed telephone lines. It is also the basis of the largest worldwide network, the *Internet.*

There is great similarity between the TCP/IP and OSI protocol suite in the lower four layers, namely, the Physical, Data Link, Network, and Transport layers. The two main protocols in the lower layers of TCP/IP suite are:

- ***Internet Protocol (IP)*** IP provides the equivalent to the OSI Connectionless Network Service. It moves data from end system to end system by means of connectionless packet delivery

- ***Transmission Control Protocol (TCP)*** TCP provides the equivalent of the OSI Transport Layer services. It establishes a connection using the connectionless packets of the IP. It therefore provides the kind of communications that programs and users are accustomed to having.

At the Transport layer, however, the models begin to diverge. Unlike the OSI model, TCP/IP does not use a Session or Presentation layer. Instead, TCP/IP application protocols each have their own *implicit* session and presentation mechanisms, and thus interact directly with the TCP, or transport layer. In OSI, the Session or Presentation layer exist as layers separate from the applications. In the end, while the services offered by the lower layers of the two models are similar, the services needed by the application layers are very different (Rose and Cass, 1987). TCP/IP application protocols, include the File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP), and the TELNET remote login protocol. All these are described in detail below.

2.2.4.3 DECNET

DECNET is an implementation of the proprietary suite of networking protocols, Digital Network Architecture (DNA), developed by Digital Equipment Corporation to enable networking of its computer systems. DNA is similar in intention and concept to the OSI suite, but because it was developed prior to OSI, networks using it are incompatible with OSI networks (Lauck *et al*, 1986).

Digital Equipment has a strong commitment to OSI and has announced its intention to ensure that future versions of DECNET will be fully compatible with OSI. DECNET Phase V, for example, is largely compatible with OSI.

There are a number of very large networks based on DECNET, such as HEPnet. Very many networks support DECNET protocols together with other protocol suites such as TCP/IP.

2.2.4.4 Coloured Book

These are the sets of protocols developed for use by JANET (Joint Academic Network), the major U.K. research network, and widely adopted by the British research community. They are also used in Ireland as well as in Australia and New Zealand (Spratt, 1986).

Although they are based on the OSI Reference Model, many of the protocols predate the full development of the OSI protocols, particularly at the upper layers. Applications developed on the basis of the Coloured Book protocols will generally be incompatible with OSI applications.

There is a commitment by the JANET community to migrate to OSI protocols, and one of the Coloured Books, the White Book, contains the migration plan (ACOSITG, 1987). Because the Coloured Book protocols are already based on the OSI Reference Model, the transition to OSI should be significantly less complex than for communities using other protocol suites (Rose, 1990).

2.2.4.5 SNA

System Network Architecture (SNA) is a proprietary suite of protocols developed by IBM for use in its systems. SNA, introduced in 1974, was one of the first network architecture to use a layered structure, and has been very influential both in its model and in specific protocols. Examples include Synchronous Data Link Control (SDLC), which forms the basis of the ISO High-level Data Link Control (HDLC) protocol, and the Link Access Procedure (LAP) of X.25 (Wheeler & Ganek, 1988).

SNA has developed in an evolutionary manner, from a simple hierarchical terminal network architecture which supported only synchronous (i.e. "online") applications, to complex internetwork architectures which support both synchronous and asynchronous exchange of information (Meijers, 1988).

Although IBM is making OSI products available, it remains committed to SNA as its strategic networking architecture. There is a very large and growing installed base of SNA networks, so much so that SNA will continue to coexist with OSI for many years to come. There are therefore major interworking issues which need to be resolved.

One very large network, BITNET, uses IBM protocols, particularly the Network Job Entry (NJE) protocol, as the basis for a store-and-forward network. NJE is not, however, formally part of SNA and it has now been implemented on a wide range of computer systems. Many BITNET sites have now changed to BITNET II, which provides BITNET services and applications over a TCP/IP network. This change has been largely invisible to BITNET users.

2.2.4.6 Other notable protocols

There are a number of other protocols suites that are used to support connectivity between computer systems. Although they are used less widely, they are worth noting.

Xerox Network Services (XNS) XNS is a proprietary suite of protocols developed by Xerox Corporation for use with its systems. It is used within the large Xerox Internet and also by numerous LANs. XNS concepts have been very influential in other architectures.

MAP/TOP Both the Manufacturing Automation Protocol (MAP) and the Technical Office Protocol (TOP) use OSI upper layer protocols on LANs. The major difference between the two is that MAP uses a token bus LAN topology while TOP uses Ethernet or Token Ring.

Network Computing Architecture (NCA) Although NCA is specified by Apollo Computer Inc., it is intended to promote resource sharing on a large scale and aims to support differing systems from multiple vendors.

Sydney UNIX network (SUN-III) The current version of SUN-III is a layered protocol suite that provides a message delivery service to support application protocols such as file transfer, electronic mail, and remote printing.

Program Support Communications Network (PSCN) The PSCN is a protocol suite which serves as the infrastructure for a number of specialized scientific networks in the United States. It was developed from the dedicated links used for these networks and uses, unlike most other network protocol suites described here, a circuit-switched network (Quarterman, 1990).

2.2.5 Protocols for Specific Applications

The purpose of application protocols is to enable distinct systems, possibly from different vendors and in different operating environments, to interoperate for a specific purpose, such as electronic mail or file transfer. An application protocol will define the kinds of messages available for transmission; the order, if any, in which they are to be sent, the data elements which

must or may be present in any given type of message, and any structure that a particular data element must possess. The main categories of application layer protocols are described below.

2.2.5.1 Electronic mail

Electronic mail (or *email*) enables individual users to send a message to one or more other users and is the most common application used on large networks and internetworks. The application is normally based on a store-and-forward model; there does not need to be a connection between the sender and the receiver at the time the mail is sent. Instead, the sender mails a message and, if the receiver is not available, it is kept in a *mailbox* for delivery at a time the receiver does become available. This mailbox may be located in a centralized location for a terminal network, or in a mail system near the receiver in a widely distributed network.

Because of the widespread demand for electronic mail, this is one of the few applications for which gateways have been developed to allow the delivery of messages between networks with incompatible protocol suites. It is therefore possible to send mail to almost anyone who has an electronic mailbox, as long as their address is known.

The creation of a gateway for mail will often involve a reduction in services offered. For instance, one protocol may allow the sender of a message to be notified when mail has been read by the recipient, while another protocol may not support such an option. The creators of the gateway need to devise strategies to deal with such incompatibilities.

Another important point to note concerning the use of gateways is that, while within a network it is often not necessary to give routing information as part of the address, mail to an incompatible network will generally need to be sent explicitly via the gateway address.

A brief description of some of the most common electronic mail protocols is presented below.

SMTP The Simple Mail Transfer Protocol (SMTP) is the TCP/IP protocol for electronic mail, defined in a document called "RFC 821". This defines an envelope to be used in delivering mail messages, together with commands and conventions for performing the delivery. The protocol requires that mail messages be in a specific format defined in a document called "RFC 822". This format uses a simple mail format, with the message represented entirely as lines of text, using 7-bit US-ASCII encoding. The basic format uses a set of *headers*, containing data such as sender's name, recipients' name, subject, and date, and a *body* which contains the text of the message.

Because the format of the mail message is entirely text-based, transfer of binary data is not possible, nor is support available for any character set other than the basic English alphabet. Thus, the protocol is awkward to use with foreign languages, and impossible for non-Roman scripts such as Japanese or Cyrillic. Neither is it possible to enclose binary data as all or part of a message, such as images, documents embedded with word processing code, or spreadsheet files.

Nevertheless, RFC 822 is the most widely used and widely copied of the mail format protocols and has had a profound influence on other mail protocols and systems.

BITNET BSMTP This is an adaptation of SMTP for the BITNET environment. It places a standard RFC 822 message, including headers, within an envelope which has addressing and control information understood by the network layer protocols used by BITNET.

Grey Book The mail format and protocol used by JANET and related networks is specified in the Grey Book. Both the protocol and the message format are similar to those of SMTP and RFC 822.

X.400 X.400 is the OSI electronic mail protocol. The name derives from the set of CCITT recommendations which first defined the protocol; the ISO name is Message-Oriented Text Interchange System (MOTIS). The full electronic mail service is defined in a number of separate protocols, of which the most important are the Message Transfer protocol (MT), which is used to deliver a message, and the Interpersonal Message protocol (IPM), which uses the MT protocol as the basis of a mail system. X.400 is very different from the protocols described above, and offers many more services:

- it supports binary as well as textual messages
- it allows multiple *body parts* within a message. For example, a message could consist of a textual body part, together with an image body part and an audio body part. Support for multimedia applications is thus inherent in X.400
- it supports multiple and nested distribution lists within the standard
- it supports a range of security features such as notification of delivery and guaranteed authenticity of sender
- address and other envelope details are recorded in binary form, and there is thus no definitive textual representation of this information; it can be presented to the user in the form appropriate to the application, and it can easily be obtained from other applications, such as a directory.

The CCITT X.400 recommendations were first published in 1984. The ISO standard was published simultaneously with revised CCITT recommendations in 1988. Most implementations to date have been of the 1984 standard, although implementations based on the 1988 standard are becoming available.

2.2.5.2 Remote login

Many large computer systems are based on the terminal-to-host model consisting of a central processor with dedicated terminals connected to it. The operator at the terminal keys in commands or makes menu selections to choose an application process to execute on the central processor as well as to interact with that process. All processing is done by the central processor and there is little or no intelligence present in the dedicated terminal. This is the model upon which most current library and information systems are based.

As the computers which house these systems increasingly become connected to research networks, there is a growing demand to be able to connect to them by means of terminals connected to remote computers. In the PC world, this is done by means of *terminal emulation* software, which allows the PC to be treated by the host computer as if it were a standard terminal. There are a large number of such software packages available.

With multi-user or time-sharing systems, however, the situation is somewhat more complicated. A terminal connected to the local system must also appear to be a terminal to the remote system. This is accomplished by means of a *remote login* protocol. Such protocols have two parts, a *client* part which resides in the local system to control the terminal and a *server* part which resides in the remote system and passes a stream of characters to the remote application as if it were a stream of characters coming from a hard-wired local terminal.

It must be stressed, however, that remote login does not provide anything more than a transparent connection to the remote system; the user must know how to invoke the required application on the remote system and then how to use it. For instance, a librarian can use a remote login protocol to connect to a remote library catalogue over a network, but unless the user already knows how to use the catalogue, or has instructions available, he or she will not be able to find any items in those catalogues. The remote logon application protocols for the major protocols suites are outlined below.

TELNET TELNET is the standard TCP/IP remote login protocol. The protocol defines a Network Virtual Terminal to provide a uniform format for transmission of data over a network. The user invokes a TELNET client application which translates between the format of the data as output by his terminal and the format as transmitted by a Network Virtual Terminal, while the remote system has a TELNET server application which translates between the format of the Network Virtual Terminal data and the format required by the host application.

The format used for transmission over the network is very simple, consisting of lines of US-ASCII text, including control characters and escape sequences. The protocol therefore does little more than pass characters from a terminal to the remote application and back again. This raises potential problems of interoperability between the application and the terminal. For example, the application might be designed to be used by dedicated, "hard-wired" terminals and will therefore not provide mechanisms for starting and stopping communication sessions

gracefully. The application may assume the use of a single terminal type and be highly tailored to that, in terms of both screen manipulation and messaging.

rlogin The rlogin protocol was developed for UNIX systems and provides remote login in a convenient manner between UNIX machines. It handles such details as identification of terminal type automatically.

Triple-X Triple-X refers to a set of three CCITT recommendations which permit remote login over X.25 networks. The protocols are as follows:

- CCITT Recommendation X.3, Packet Assembly/Disassembly (PAD), specifies how character strings emerging from a terminal are to be converted into X.25 packets and vice versa. It describes the functions that are performed by the PAD facility and defines the various parameters that can be used to specify its mode of operation

- CCITT Recommendation X.28 defines the interface between the PAD and the terminal. It describes the way a terminal connects to and controls the PAD facility

- CCITT Recommendation X.29 defines the interface between the PAD and the X.25 network. It describes the way the X.25 host communicates with the terminal by way of the PAD. This protocol makes use of a Virtual Network Terminal, similar to that of TELNET.

X.3 and X.28 are normally implemented in the PAD while X.29 needs to be implemented in both the PAD and the host computer.

It should be noted that the Triple-X protocols were standardized very early and do not adhere to the OSI Reference Model. In particular, they do not use the upper layers of the model; rather they make use of the X.25 network layer services directly.

Green Book The JANET remote login protocol is specified in the Green Book. It specifies how the ISO Triple-X protocols are to be used in the JANET environment.

Virtual Terminal Virtual Terminal (VT) is a more recent OSI protocol to support remote login. It conforms fully to the OSI Reference Model, and offers considerably more sophistication than Triple-X, particularly in its ability to define "profiles" for commonly encountered terminal types and the ability to negotiate support for terminal characteristics.

The protocol is now an international standard and commercial implementations are becoming available.

X Window and Motif X Window is a distributed, network-transparent, device independent, multitasking windowing and graphics system. Multiple applications can be displayed on the same screen, each within their own window. X Window also allows a single application to use

many windows. Windows can overlap or be hidden. The system also supports soft fonts and two-dimensional graphics drawing.

X Window achieves its distributed, network-transparent and device independent characteristics by splitting its functionality into two parts using the client/server model. The client is an application program which makes requests of the server program to draw windows, text and other objects. The server is a program running on each workstation which draws the requested objects on the display. Each workstation has its own server which contains hardware dependent device drivers for that workstation.

The client sends packets of instructions to the server which conform to the X Protocol, which is high-level graphics-description language. The client and the server may be resident on the same workstation, when a single user is executing a program locally; or they may reside on different computers communicating with one another over a network.

X Window provides the mechanism for drawing windows and sending graphics over a network, however, it does not define the user interface that runs on top of it. The user interface specifies what windows and icons look like, how the user interacts with them, and so on. Motif is a user interface based upon X Window which is emerging as a GUI (Graphical User Interface) standard, particularly for Unix systems.

Motif was developed by the Open Software Foundation, which includes such vendors as IBM, HP, DEC and more than 40 others. It is designed to be totally device independent. It is currently available for a variety of Unix-based platforms.

2.2.5.3 File Transfer

File transfer refers to the ability to transfer files between remote hosts and possibly manipulate remote files, such as deleting them or changing their attributes. Although it is possible for files to be transferred using electronic mail, this is unsatisfactory for anything but the most simple file transfer. Difficulties arise because: 1) the file will usually need to be manipulated to place it into a message, 2) many protocols will allow only text files to be included in a message, and 3) only short files are appropriate for this kind of transfer. File transfer, on the other hand, offers the possibility of direct transfer of files of almost any length and with any structure.

The following describes some of the most widespread file transfer protocols.

FTP The File Transfer Protocol is the TCP/IP protocol for file transfer. It is based on peer-to-peer interactions between a local and a remote host. As with many other TCP/IP protocols, it supports only a basic set of functions:

- it supports only the transfer of files

- only complete files can be transferred
- only a few types of file structure are explicitly supported: binary files, ASCII text files, EBCDIC text files, and "paged" files.

One of the more common uses of FTP is the convention for "anonymous FTP", which allows a user to retrieve selected files from a remote system without prior authorization. This allows for rapid and informal publication of information across a network, without needing to mail potentially long documents to many users. Both text and program files are frequently distributed in this way.

Blue Book The JANET file transfer protocol is currently the Network-independent File Transfer Protocol (NFTP), and is specified in the Blue Book.

FTAM The OSI protocol for file transfer is File Transfer, Access and Management (FTAM). The protocol provides a full file service, including open, close, read, write, delete. It offers access to the whole file and also to parts of file. In structured files, it offers access to single or multiple records. It offers access control, file and record locking, accounting and error recovery features.

Note, however, that FTAM is not a database; no indexing or searching (other than on record keys) is supported, and only a single file may be open at a time. This latter restriction means it cannot easily be used to provide a distributed file system for applications that require many files to be open together, such as database look-up, which requires concurrent access to multiple index files and database files.

An alternative to using a file transfer protocol is to provide network-wide file operations similar to those available from the local operating system. In this approach, the distinction between local and remote files is invisible to the user. The best known of these protocols is the Sun Network File System, which has been widely implemented on Unix systems.

2.2.5.4 Remote Job Entry

Job entry is the submission of a set of instructions for one or more programs to be executed as a batch. It is one of the oldest types of interaction with a computer system and remains of considerable importance in computer systems of all sizes. Remote job entry is therefore an important application for networks.

However, job entry itself and its management is normally a relatively specialized task, often conducted from a computer centre, so only a brief outline of the major protocols will be given.

Job Transfer and Manipulation (JTM) is the ISO protocol for remote job entry. It allows for the movement of both input and output data across a network as well as movement of monitoring and control information.

The equivalent JANET protocol is specified in the Red Book.

2.2.5.5 Directory Services

The effective use of network services depends on the existence of directories, just as telephone service depends on telephone directories to allow users to find the correct number to dial. Similarly, as telephone directories can be either formal ones, such as the White Pages published by the telephone companies, or informal ones kept personally in notebooks or address books, so can network directories be formal ones publicly accessible via the network or informal ones kept locally for local use.

The aim of a network directory service is to allow network users, both people and other computer applications, to look up network identifiers, addresses, or other information about other users in the same way that a printed telephone book allows telephone numbers found.

The simplest form of directory service is a fixed or "static" table held by each host machine in a network which gives the address and routing instructions for every other host reachable on the network. A network, however, is seldom static for very long; new machines are added, old ones are removed, and communication links are altered. The task of keeping a large number of fixed tables up to date therefore rapidly becomes impossible as a network grows in size.

All the major protocol suites include protocols for some kind of distributed directory service. Most of these are very simple, limited to providing a mapping from a host name in the format required by the protocol to the network address. It is possible, however, as the discussion of the OSI Directory will illustrate, to conceive of much more ambitious directory services.

Non-OSI protocols for directory service include:

DNS The Domain Name System is the TCP/IP directory protocol. It implements a hierarchical tree-structured naming scheme and provides protocols for communicating with *nameservers* on the network which map names to TCP/IP addresses.

TCP/IP addresses consist of 4 8-bit bytes, usually written as 4 decimal numbers separated with periods: *225.15.22.8*. This address identifies a TCP/IP host. Individual users are not explicitly identified in addresses.

The DNS name scheme allows a name to be assigned for each address in a hierarchical manner using a tree structure as illustrated in (see Figure 2). At the top level of the tree, a number of naming "domains" have been identified. These are individual countries, identified by a two letter country code such as "ca" for Canada, "uk" for the United Kingdom, and "fr" for France.

The United States, however, does not have a single domain, but has several domains at the top level, such as "com" for commercial organizations, and "edu" for educational institutions. These top-level domains are assigned by the *Internet* Advisory Board.

Each domain is responsible for assigning names at the next level in the hierarchy. Similarly, each sub-domain is responsible for assigning names at the next level, and so on, until individual hosts have names consisting of a hierarchy of domain names: *cunyvm.cuny.edu* names a host (*cunyvm*) in a subdomain (*cuny* = City University of New York) of a domain (*edu* = US educational institutions). Because of the tree structure, names need only be unique at each level within an individual domain. Thus *cuny.com* would name a different host from *cuny.edu*. This allows the assignment of names to be highly decentralized.

These domain names do not correspond to particular networks or subnetworks; hosts in the same domain may be on different networks, and hosts in different domains may be on the same network.

In order to discover the address corresponding to any given domain a network user makes enquiries at one or more nameservers. There are a number of nameservers at the *root* level, from which the addresses of nameservers at the next domain level can be determined until, after a series of messages, the address of any host can be discovered. To reduce the number of messages needed, nameservers keep full name to address mappings for names which have been queried recently. Thus it is possible that the address of *cunyvm.cuny.edu* will be obtained from the root nameserver. This is known as *name caching*.

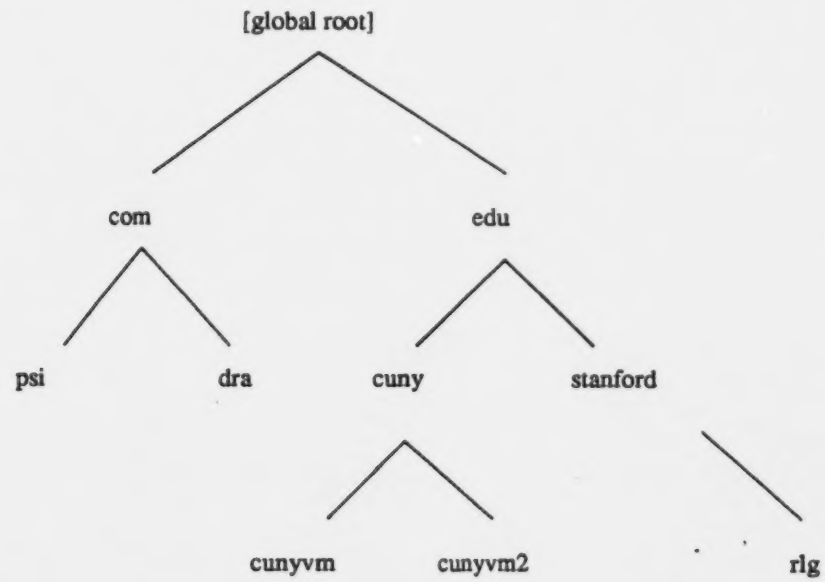

***Figure 2** - Hierarchical Tree Structure of Domain Name System*

Since user names are not included as part of the DNS, applications such as electronic mail which require them must have a means of distinguishing a user name from a host name. This is done by means of the "@" convention. A user name is separated from the host name in an electronic mail address by means of an "@" sign (read "at"). For instance *smith@cunyvm.cuny.edu* addresses a mailbox named "smith" at the host *cunyvm.cuny.edu*.

Many networking communities using other protocol suites have agreed to use the DNS naming conventions and domains as a common naming syntax.

JANET DNS The JANET Name Registration Scheme (NRS) is equivalent to the *Internet* DNS. The NRS supports *campus name servers* with the *Name Lookup Protocol*, allowing different addresses for different application protocols and services on the same host, and provides the essential basis for the transition to OSI protocols (Stone, 1991). The naming conventions use hierarchical domains similar to DNS, but the domains are listed in the opposite order. The name *uk.ac.ucl.cs* for instance identifies the Computer Science department at University College London. The NRS also permits registered abbreviations to be used, so that the name *manchester.central-services.pacx-b* can also be addressed as *man.cn.xb*. Note that within JANET the domain prefix *uk.ac* is often omitted.

X.500 Directory service

X.500 refers to the set of CCITT Recommendations for the OSI Directory Service. The counterpart ISO/IEC standard is ISO/IEC 9594. The Directory defines a distributed database for OSI applications. It contains information about people, applications, and devices, and provides structured mechanisms for accessing the information. This information can then be used by both persons and other OSI applications.

One of the primary goals of the Directory is to enable "user-friendly" naming in the OSI environment. This will permit a user to specify a name, such as "laser printer", "Department Head", "Phil Jones", and then to retrieve additional addressing and possibly other information which can be used by either the user or an application. Thus a telephone number might be retrieved for a user or a network address by an application. The OSI Directory is thus much more ambitious than the TCP/IP or JANET services described above.

The Directory is not, however, intended to be a general purpose database, and it cannot be expected to take the place of generic database applications such as RDA or SR. In particular, the Directory does not guarantee the consistency of its data, and short-lived errors in consistency, such as duplication of data, may be common.

The Directory comprises a single distributed database, known as the Directory Information Base, which may be interrogated and modified by a user. All access is subject to both authentication of the user and authorization by the service provider.

The Directory uses a hierarchical tree naming system similar in principle to that used by the TCP/IP DNS as described above. It is however, very different in detail. In particular, the structures used by the Directory are much more generalized than those of DNS. While DNS deals only with network resources such as hosts and routers, entries in the Directory can be made for almost anything including people, machines, programs, and organizations.

To accomplish this, the Directory Information Tree consists of *entries*, each of which describes, by means of various *attributes*, a single *object*. The object belongs to a *class* which defines what types of attribute it must or may have.

Entries in the Directory Information Tree are arranged in a tree structure in which groups of entries have a single superior entry to form a subtree and groups of these subtrees in turn have a single superior entry until finally a single root is reached.

Entries are identified by a name, which can be defined differently for each object class, and may be taken from a single attribute or from more than one. Entries at the same level in a subtree have a ***relative distinguished name*** which is unique. This relative distinguished name need not, however, be unique within the whole Directory since each entry can be uniquely identified by its full ***distinguished name***. The distinguished name is comprised of its relative distinguished name together with the relative distinguished name of each of the entries above it in the hierarchy.

Thus, for an employee of a library, the Distinguished Name (DN) might be:

countryName=Canada
locality=Ontario
locality=Ottawa
organization=National Library of Canada
organizationalUnit=Information Technology Services
organizationalUnit=User Support
organizationalPerson=A N Other

Equally, the Directory can be used to identify computer systems or applications:

countryName=Canada
locality=Ontario
locality=Ottawa
organization=National Library of Canada
applicationProcess=PEB/ILL
applicationEntity=ILL module

The Directory does not specify the relationship between objects in the hierarchy. This kind of specification is dependent on the purposes for which the Directory is to be used, and it can easily be envisaged that different users of the Directory would want different rules. The Directory is therefore seen to consist of a number *sub-trees*, each of which can have different structure rules

that specify the relationship of object classes at different levels. For instance, one subtree may require that persons are always subordinate to a locality (i.e. a person lives in a town), while another may require that persons are always subordinate to organizations (a person works for a company). Yet another subtree may allow a person to be subordinate to either.

The combination of the various sets of rules determining how object classes are related, such as the attributes that comprise an object, and the syntax of those attributes, is known as a *Directory Schema.*

The Directory may be interrogated in several ways and may also be updated by those with the appropriate authorization.

Commercially available implementations of the Directory are beginning to appear, and large-scale pilot projects are taking place, notably the "White Pages" project sponsored by NYSERNet (see 4.1.1).

2.2.5.6 OSI library-oriented protocols

This final section of the overview of networking describes two OSI protocols intended specifically for library applications. They have slightly different sources than the more general OSI protocols. International standards relating to libraries originate with the ISO Technical Committee for Documentation (TC46). Neither the IEC nor CCITT have any involvement with this work, which is undertaken primarily by librarians and those interested in library systems. The parallel American organization is the National Information Standards Organization (NISO), an ANSI-accredited organization responsible for standards in the area of libraries, publishing, library equipment, and information services.

Z39.50 and SR

Z39.50 and SR are parallel information retrieval protocols; Z39.50 is the NISO standard, while ISO Search and Retrieve (SR) the ISO standard. Z39.50 and SR are functionally equivalent, with the exception that Z39.50 supports access control and resource control while SR does not, and the 2nd version of Z39.50 is now expected to be compatible with SR. They will therefore be treated here as a single standard.

The purpose of the standard is to allow searching of databases on a remote system and retrieval of records as the result. To accomplish this, the standard posits a simple model in which an *origin* system transfers requests to a *target* system and receives replies from it.

The target system is assumed to have one or more *databases*. A database consists of at least one and possibly multiple files which contain a single set of *records*. All the records within a given database have one common structure, and may be retrieved on the basis of a common set of *access points*.

An access point may or may not identify a record uniquely; and it may be equivalent to a data element in the record structure or may be derived from a data element or may be the result of combining all or part of two or more data elements.

Note that the standard does not specify the structure of a database, the structure of a record, nor the structure of an access point. In particular, the standard assumes neither the presence nor the absence of indexes. Finally, the standard does not specify the types of search term a database is to support, in either content or structure.

To retrieve records in the database the origin sends a search request to the target. This request contains a query and specifies such things as the ID of the database to be searched and the action to take in the event that the search retrieves a small, medium or large set of records.

The standard defines three types of query: 1) a private query whose nature is agreed bilaterally between communicating parties, and is not further defined, 2) a query structured according to ISO 8777, the Common Command Language, and finally, 3) the *RPNQuery*, which is the only one that conforming implementations need support.

The RPNQuery uses a Reverse Polish Notation to structure complex queries. This notation uses the structure "*term term operator*" (e.g. the addition of the numbers one and two would be written: "1 2 +") It has the advantage over a command language that it is very easy for a computer processor to parse, and because the result of each operation becomes the first term for the next, allows indefinitely complex queries to be created.

Each search term in a query has associated with it one or more *attributes* which define its nature and structure. These attributes will often define an index type or operation in the target database. For instance, the term "Gulliver's travels" might have the attributes "title" and "phrase" associated with it, specifying that the search term is the title phrase "Gulliver's travels".

The attributes which may be associated with a term are not, however, defined as part of the standard, although one set of attributes, the "Bib-1" attribute set, is included as an annex. This set of attributes is useful primarily for searching of bibliographic databases. Other attribute sets are likely to be developed for other purposes such as searching full-text databases and audio-visual files. To enable origins and targets to interoperate, attribute sets are registered with either national or international standards bodies, and are intended to be widely published.

The target sends a response to each search request, indicating the success or failure of the search, the reason for a failure, and possibly a set of records. The target may also retain the result of a successful search as a *result set*.

Whenever a result set exists, the origin may request one or more records from the result set to be "presented" to it. It may specify the number of records to be presented, the preferred record composition, and the starting point in the result set.

A result set is normally assumed to be available for reference from the time it is created until the search session ends. It is assumed to remain available even if subsequent searches have taken place. The origin may therefore request that a result set should be deleted.

There is considerable interest in the use of Z39.50 and SR to provide access to remote databases. It holds the promise of allowing the creation of systems which provide a single coherent user interface to multiple databases, thereby reducing training costs, and making information available to a much wider audience than has previously been possible.

The ILL Protocol

This protocol is an ISO standard for interlibrary loan applications. It can also be used for many, but not all, distributed circulation functions. It supports a number of different interlending structures, from the simple model of one library approaching another library to borrow an item, to complex ILL networks involving centralized location agencies which undertake request routing, or to decentralized structures in which requests may be chained from one library to the next until the required item is supplied.

The basic model of the ILL protocol is of a "requesting" library, which wishes to borrow an item, a "responding" library, to whom the request is sent, and possibly one or more "intermediaries".

The protocol is transaction based, with each complete instance of a whole interlending cycle forming a single transaction. Multiple concurrent ILL transactions are allowed for any party in the ILL process.

The protocol provides a large number of message types related to the negotiation and supply of an item on ILL and its subsequent tracking during the period of the loan. Each of these message types is normally provided by only one of the parties in an ILL transaction and is not normally acknowledged.

One significant feature of ILL is that it does not use a MARC-structured record to provide the item identification. Instead a simple set of data elements to allow item identification is defined within the protocol.

The ILL protocol was recently approved as an International Standard (IS) in April, 1991.

3. LIBRARY APPLICATIONS OF RESEARCH NETWORKS

Many vivid phrases have been coined to suggest the uses to which research networks will be put in the coming decades, such as *information superhighways* (Gore, 1990), *colibratories* (Larsen, 1990), and *virtual libraries* (DeCandido and Rogers, 1990). These phrases evoke a world in which students, scholars, and researchers access vast amounts of information directly from workstations located in their homes and offices. Hidden beneath these broad images, however, is the fact that such functionality will require the development of a number of specific applications to perform the various tasks involved. Recall that research networks are essentially general communication infrastructures that can be used to accomplish any number of tasks between remote systems and users. They allow low-level, peer-to-peer communication between computers, but the specific tasks they support are accomplished through higher-level services, or applications—the actual services that are provided to end-users, be they human or machine. Examples of common network end-user applications are electronic mail and file transfer. Within library science, specific end-user applications of research networks fall into the traditional domains that have characterized librarianship. These domains are *bibliographic information provision*, involving a library's mission to mediate access to information, *document supply* involving supplying the actual document through interlibrary loan or other means, and *technical services* involving the behind-the-scenes tasks, such as cataloguing and acquisitions, that indirectly support information provision and document supply.

Communication among end-user applications is defined by *application protocols* which are specific statements as to the rules and formats that communicating applications must adhere to to perform the desired function. Application protocols for electronic mail, for instance, are SMTP on TCP/IP networks and X.400 on OSI networks. The application protocols that can be used to support library applications are those that are general to networking, like X.400 and FTAM, as well as those specifically developed for libraries, such as Z39.50 (or SR) and the ILL protocol. Other standards that can be employed for library applications include ODA (Office Document Architecture) and EDI (Electronic Data Interchange). While protocols in other protocol suites have been used for library purposes, for example, TELNET access to *Internet*-accessible catalogues, this section will focus on those protocols that are based on the OSI reference model.

By employing a number of application protocols and standards, libraries can create end-user applications that greatly enhance access to information, document supply, as well as the efficiency of technical processes. Within a network environment, for example, local bibliographic information provision can be augmented to allow distributed information retrieval of multiple remote databases. Document supply can also be improved, accelerating the supply of documents by sending them directly to the end-user in electronic form. For technical services,

research networks offer direct machine-to-machine exchange of cataloguing records, authority records, as well as acquisition information.

In all three domains, research networks have the potential to bring a contemporary library's use of automation and telecommunications to a higher level of functionality, and offer a method by which disparate processes dependent upon such techniques can be rationalized. This chapter will discuss specific areas within these domains that have greatest potential for research network applications, outlining the impetus for each area's use of networks, and the application protocols that are most appropriate for supporting the applications.

3.1 Applications for bibliographic information provision

The first traditional library domain in which research networks can be utilized is bibliographic information provision, consisting of the public, user-oriented tasks involved in helping end-users find information relevant to their work. These tasks traditionally include assisting users in finding relevant citations through information retrieval, reference service, or selective dissemination of information. The vast amounts of information that can potentially be accessed by a rapidly growing number of network users has been called the real value of research networks (Cline, 1990; Larsen, 1990). The areas of bibliographic information provision that have greatest potential for use over research networks are information retrieval and the selective dissemination of information (SDI).

3.1.1 Information retrieval

Information retrieval involves the process of searching a database for relevant materials on any given topic, and thus has become an important tool in support of the library's function to mediate access to information. Since the development of early online information systems and public-access catalogues in the 1970's, systems supporting information retrieval have proliferated (Arms, 1989b). Search software has become increasingly sophisticated, supporting techniques such as boolean operators, keywords, wild cards, multiple access points, and inverted indexes. Databases available for searching currently number in the hundreds, covering almost every topic and academic discipline, and exist in a variety of formats from bibliographic to full text. The power of IR systems to find information quickly and efficiently, and the large number of electronic databases currently available, has made this technology indispensable to both libraries and their patrons.

Information retrieval is also an integral component of the electronic libraries to be supported by research networks. Most discussions about the ideal electronic library focus on the ability to retrieve information easily from a variety of sources. Various conceptualizations of networked electronic libraries have essentially three common characteristics (from Kibbey and Evans, 1989):

- ***location independence***, or the ability to access information from the workplace, regardless of where that workplace is located

- ***breadth of contents***, or the ability to access primary information in multiple databases of various formats

- ***ease of use***, or the ability to use the electronic library without detailed knowledge of various interfaces, search conventions, or database formats.

In contrast to this ideal, present-day libraries have been forced to develop a collection of disjunctive strategies to supply information to end-users. Information retrieval in the majority of libraries is characterized by:

- ***many distinct resources***. Libraries have had to cope with a massive increase in the quantity of recorded information and in the systems developed to manage it (Arms, 1989b). Most libraries, for example, possess online public-access catalogues that provide access to their own collections. These catalogues are limited, however, to primarily monographic material and to those items actually housed within the library. While access to materials beyond the local catalogue have always been available in book form (Jacob, 1990), online catalogues of remote libraries typically have been unavailable to local users. Also unavailable in online catalogues is access to the rich body of journal literature that forms the basis of most scholarly communication (Lynch, 1989a). Such information is contained in the multitude of separate bibliographic, abstract, full text, image, and numeric databases that have been created for most subjects and academic disciplines. Libraries must provide access to these resources through any number of commercial and other information services. To fully satisfy the information needs of their patrons, therefore, libraries have been forced to provide access to this large collection of disparate information resources and systems

- ***diverse user interfaces***. Different databases, database vendors, and online catalogues typically have diverse user interfaces. They can vary in screen layout, overall functions provided (e.g. keyword searching, browsing), command languages, order of operator processing, stop words, and in many other areas. This fact necessitates both learning new interfaces every time another system is searched and differentiating between details of systems that are used regularly. This is a major problem if an assortment of databases and database hosts are utilized. Search interfaces and conventions are so varied and complex that searching online database is still largely done by expert intermediaries

- ***multiple communications channels***. Many of the systems must be accessed through different channels, for example, a dedicated terminal in a library, a dial-up line to any number of information vendors, or a connection to a bibliographic utility's private network. For libraries, having separate communications channels results in having to acquire and maintain additional equipment, pay for numerous accounts with PDNs, and train both staff and users in the use of the various modes

- ***limited end-user access.*** Those who actually use the information, the end-users, have limited direct access to the array of information resources maintained by a library. This manifests itself in two ways: 1) limited access to *end-user* searching and 2) limited *physical* access to systems that support information retrieval. Until very recently, the only avenue for end-user searching was the public-access online catalogue (Arms, 1989a). Other types of databases, like those available from commercial information services, remained the domain of information specialists because of the high costs involved. Experiments with providing end-user searching of databases other than the catalogue, however, have revealed a staggering demand for these facilities (Lynch, 1989a). Some major university libraries have begun to respond to this demand by promoting end-user searching though user-friendly interfaces to remote online databases, installing CD-ROM systems, or mounting databases for access on local machines (Arms, 1990a; 1990b). Yet, to genuinely make information universally available, users should be able to have physical access to information retrieval facilities wherever they are working. Physical access to information is limited if students and researchers must come into the local library to search various forms of databases, or go to a library out of the area to search its catalogue.

In the end, the current approach that contemporary libraries must take to provide access to a full range of information resources requires strategies that involve multiple information resources, disparate user interfaces, and numerous communications channels. Further, users are typically isolated in their work places from these resources. Research networks have the potential to unify these strategies by providing a single, widespread communications infrastructure over which information resources can be made available. Library applications, based on application protocols such as Z39.50, could put this diverse set of information resources into a unified framework that would be directly available to the end-user thereby providing greater access to information.

Use of the research network infrastructure

A number of academic libraries have begun to make a variety of information resources more widely accessible to their users through research networks. The first step for many institutions is to connect library systems to campus-wide networks (see Kinney, 1988; Arms, 1990a; Lynch and Preston, 1990). While a full discussion of campus-wide networks is outside the bounds of this paper, these networks are basically high-speed LANs or MANs that span university or college campuses, interlinking PCs, workstations, mainframes, printers and other peripherals. To this network, libraries connect information resources such as online catalogues and locally mounted journal databases. The connection provides end-users with access to these resources from anywhere they can access the campus-wide network: offices, laboratories, dormitories, and homes. Campus-wide networks are usually linked to mid-level research networks which are in turn linked to larger networks. These interconnections make the resources available over a widespread area that can be international in scope.

It is in this way that a large number of online catalogues have become available over national research networks. There are, for example, probably a total of over 260 catalogues connected to research networks as of the spring of 1991 (Lynch, 1991; Stone, 1991). Users can search these

catalogues located at libraries across the country, or on another continent, from their local workstations. Such widespread access begins to meet the criteria of location independence of the electronic library.

These initial attempts, however, fall short of the other two criteria, that is, ease of use and breadth of contents. To use the online catalogues linked to the *Internet* and JANET, for example, one must first know the proper TELNET or NRS address, the login procedures and passwords, as well as the idiosyncrasies of the interface of the particular library system to which one has connected (see Chapter 4 for a full discussion of *Internet*-accessible and JANET-accessible library catalogues). Such user-unfriendliness restricts the ease of use of the networks for remote searching. Further, because the resources are located on distinct systems, the results obtained from them cannot be combined; searches would have to be re-submitted for each system, with the multiple result sets potentially containing many redundant citations. Breadth of contents is limited by this diverse set of interfaces, as well, in that learning a new interface for each system searched would act as a disincentive to search numerous databases. A further factor limiting breadth of contents is that most locally mounted bibliographic databases are restricted to only a select few by licensing agreements.

Distributed information retrieval

What is required to provide breadth of contents and ease of use is a way to access a variety of physically separated databases with a consistent user interface. To accomplish this type of *distributed information retrieval*, the bond between a specific search interface and its database must be broken; that is, one search interface must have the ability to do searches on any database, regardless of the specific interface that was developed for that database. Research networks, because they are peer-to-peer networks, allow the separation between workstation and database by changing the *master-slave* or *master-servant* relationships common to most online systems, to *client-server* relationships (Arms, 1989b).

Master-slave relationships are common in terminal-to-host networks, such as the dedicated terminals within libraries connected to turnkey systems; the user interface that the user sees, the software that manipulates the data, and the data itself, are all located in the central host. Master-servant relationships, on the other hand, locate many interface management functions in the local workstation. This type of terminal-to-host relationship is typical of commercial information services in that the local machine manages most of the local activity while searching is handled by the remote host. In both of these configurations, however, the software in the local terminal is bound completely to the remote database; that is, the commands and search conventions are database-dependent. Search software designed for use with a BRS host, for example, cannot be used with a DIALOG host, or vice versa.

The bond between a specific search interface and its database is broken in client-server relationships. This approach is possible in peer-to-peer networks in which each connected system is considered to be equal to any other system regardless of type, size, or operating system (Arms, 1989b). There are no longer masters, slaves, or servants, but equal computers that act, when necessary, as clients and servers. The server provides a specific service to the network which any

client on the network can use (Arms, 1988a). In this way, any one machine can communicate with and use the resources of any other machine on the network. For information retrieval, the host holding a database becomes a *database server*, supporting the database and a search engine, which the client application can access. The client application is thereby database-independent, able to search any database at any location as if it were its own.

The Z39.50 protocol The actual format and sequence of messages sent between a client and a database or information server is defined in the NISO Z39.50 application protocol, and in its ISO functional equivalent, SR (see 2.2.5.6 for more detailed information). This protocol originated with the Linked Systems Project as a way to exchange bibliographic information between utilities, but has since been expanded to include links between all bibliographic systems. The protocol specifies how, and in what form, search requests and results will be managed between the client and server (Buckland and Lynch, 1988). It does not include a definition of the actual command language or user interface, but the underlying functions and the machine-to-machine communications. In this way, any search interface defined as a client application could formulate and forward searches to another system, retrieve the results, and display them to the users in a way that is familiar. This is accomplished through the use of translation of the local search syntax into an intermediate syntax which is then translated into the syntax used by the remote system. Note that it is the local system that forwards the searches rather than the user; the user would not necessarily have to know where the remote databases physically reside, or how many databases were accessed to find the information.

Employment

Using the Z39.50 or SR protocols, applications could be created to allow distributed information retrieval over OSI-based research networks, or networks that have adopted transition strategies. Such applications would relieve users from having to deal with the complexities of many retrieval systems, and allow them to access a variety of physically separated databases with a single user interface from their places of work. In addition, the large number of databases that currently must be accessed by users and maintained by libraries may be brought into one unified, coherent framework. Furthermore, the variety of strategies needed to support information retrieval, along with the concomitant telecommunications channels, would be rationalized in this framework. Although many technical and political problems remain to be resolved (see Chapter 5), such an application will at least make a start at fulfilling the criteria of the "ideal" electronic library of location independence, ease of use, and breadth of contents.

3.1.2 Selective Dissemination of Information (SDI)

The second library function related to the domain of bibliographic information provision that can be applied to research networks is the selective dissemination of information (SDI). This service, also known as current awareness service, provides a client with regular reports of new material on topics of interest. Depending on the level of service, this might be a list of new accessions to the library, recent journal articles, or even extracts from full text databases. SDI is normally based on a user-specific "profile" of subject topics, such as classmarks or author's names, to be searched in

specified databases. New additions to the databases are searched regularly for items matching the profiles and are then sent to the client.

In the past, SDI service has been labour intensive and expensive because it required highly trained staff to search multiple databases. The use of research networks for SDI, however, promises to make this service more economical and efficient. While there are no specific application protocols for SDI, SR or Z39.50 and electronic mail have potential for use in this area. For example:

- the use of the SR or Z39.50 protocol should minimize the need to use different search strategies for different databases and allow regular searches to be done automatically
- electronic mail could be used to deliver results of the SDI service, whether citations or full text, to the client

One potential danger involved with user-initiated SDI are the use of *search scripts*. These are small programs designed to automatically login to a database host, perform searches, and send the results back to the user's workstation. Search scripts potentially may cause hundreds of unattended searches to be performed at unpredictable times thereby congesting library systems (Lynch, 1989b).

3.2 Applications for document supply

Document supply has been defined as "the transfer of a document or a surrogate from a supplier, whether a library or a document service, to a requesting library" (Boss and McQueen, 1983:1), or more widely as "the chain of provision and supply of publications from their originators through intermediaries to their users" (Plassard and Line, 1988:1). It is a broad term that is used to describe a wide range of activities involving publishers, document supply centres, information brokers, and libraries, that have the end goal of delivering documents to users. The two major methods of supplying documents to users which research networks can potentially support are interlibrary loan and electronic document delivery.

3.2.1 Interlibrary loan

Interlibrary loan is a vital part of the resource sharing activities in which libraries engage, supplying actual documents to users that are not available locally. Specifically, the interlibrary loan process entails the mutual lending and borrowing of materials between libraries, from the exchange of messages involved in the transaction process, to delivery of the document. The following discussion will focus on the process by which ILL messages are exchanged, leaving discussion of the physical delivery of the document to a separate section.

As in other domains of library work, interlibrary loan has used computers to automate its processes and, as a result, has realized significant improvements in its service over manual methods (Turner, 1990). Automation in interlibrary loan consists of various messaging systems that facilitate the exchange of ILL messages between lending institutions. Earliest methods were teletype-based (TWX) systems which were later replaced by ILL subsystems operated by bibliographic utilities, such as OCLC, RLIN, and WLN in the U.S. (Jackson, 1990). In Britain, most requests are handled by the British Library's Document Supply Centre (BLDSC) (Stone, 1990a) through ARTTel, an online service for requesting loans or photocopies from the BLDSC near Boston Spa, West Yorkshire (Buxton, 1988).

Similar to other domains in librarianship, improvements brought by automated systems have been accompanied by conflicting telecommunications strategies to carry out these functions. In the current ILL environment, there are a variety of conflicting messaging software to create ILL messages as well as telecommunications channels to exchange the messages between libraries. These conflicting strategies result in a number of barriers to ILL communication (Turner, 1990). These barriers include:

- an inability to exchange messages between different telecommunications networks. Examples are public electronic mail systems (ON-TYME II, ALANET), subject-oriented networks, networks specific to a geographic location (ILLINET, DOCLINE), and bibliographic utilities (OCLC, RLIN). ILL messages originating on one system cannot be sent to another system, and therefore there is a restriction in the flow of messages

- conflicting ILL message formats. Message formats vary with ILL networks, bibliographic utilities, and particular ILL software. This wide variance among message formats prohibits automatic processing by dissimilar systems, regardless of whether they could exchange messages across telecommunications channels or not

- other messages, in addition to the ILL request, need to be exchanged in order to complete the entire ILL process, such as *shipped*, *received*, *renew*, *overdue*, and *recall*. The number, form, and sequence of these messages remain unspecified across ILL systems.

Because these barriers prohibit communication among disparate ILL systems, libraries are required to maintain subscriptions to more than one system to gain access to the full range of libraries that lend materials.

Use of the research network infrastructure

It is apparent that two fundamental problems need to be resolved to enable the full automation of ILL processes: 1) the standardization of ILL message formats and, 2) the use of a coherent telecommunications infrastructure for the exchange of messages. An application for ILL messaging on research networks has the potential to solve both of these problems. Protocols such as the ILL protocol, which standardizes the form, content, and sequencing of ILL messages, and the X.400 application protocol, which defines message handling systems, are available to develop

a method to exchange ILL transactions among libraries who have access to electronic mail. These protocols are outlined below.

The ILL protocol The ILL protocol is a standard based on the OSI reference model that defines a standard for "the exchange of ILL messages between bibliographic institutions that use dissimilar computers, systems, and communication services" (Turner, 1990:75; see also 2.2.5.6). The ILL protocol allows heterogeneous systems to exchange ILL messages without affecting local processing and operational requirements. It additionally formalizes and automates existing processes that were previously done manually. Aspects of the ILL transaction that the protocol formally standardizes are (Turner, 1990):

- the number and type of messages to exchanged in a transaction. That is, what messages will be exchanged in addition to the ILL request. Messages types include *ill-request, ill-answer, shipped, received, renew, renew-answer, overdue, recall, returned, checked-in, lost, damage, cancel,* and *cancel-ack*

- the data elements contained within the messages, for example, the date, time, ISBN, and ISSN

- the sequence of messages, ensuring that messages, such as *shipped* and *received*, are exchanged in a logical order

- the transfer syntax to facilitate machine-to-machine communication. This involves the specification of an encoding scheme to allow the sending and receiving computers to recognize specific parts of a message.

One aspect that the ILL protocol does not specify is the manner in which messages can be sent from one type of telecommunications channel to another. The actual delivery of the ILL messages from one institution to another can be managed by systems such as electronic mail. The application protocol for electronic mail in the OSI protocol suite, referred to more generally as message handling systems, is the X.400 protocol.

Message handling systems and ILL As discussed in section 2.2.5.1, message handling systems are electronic mail applications based on the OSI X.400 protocol that assist users in the preparation, delivery, and filing of electronic messages (Swain and Tallim, 1990). Conceptually, these systems manage the delivery of mail from place to place by providing an addressable "envelope" into which users can insert messages. Note that *users* may be either human or machine and that *messages* can be more than the personal correspondence that characterize regular electronic mail; they can also refer to any formalized, structured body of data that may be exchanged between computer applications. This support for non-interactive traffic between systems is of special significance. In most email implementations, users interact directly with the system in an online, real-time mode (Swain and Tallim, 1990). Applications using X.400 MHS, on the other hand, can automatically exchange messages in the background without user intervention. This attribute makes MHS especially useful for the exchange of

messages between specific, higher-level applications, such as ILL systems, residing at remote locations.

Employment

As research networks migrate toward OSI standards, or as interim measures are adopted, implementations based on the above protocols can be developed to support ILL processing. An implementation may be designed to work in the following fashion. Forms based on the ILL protocol can be called up, filled in with the appropriate citation information and other transaction details, and automatically forwarded to the lending institution by an X.400 MHS. A peer ILL application at a receiving institution can then accept the message, extract the appropriate data elements, and proceed with further processing. Advantages of such implementations include off-line preparation and storage of messages, reductions in clerical activities and the need to train staff to operate disparate ILL systems (Turner, 1990). Further, the large number of disparate ILL systems that need to be maintained have the potential to be rationalized within one framework, reducing costs, procedures, as well as increasing the scope of potential lending institutions.

Since research networks extend to most higher education and research institutions where the majority of research libraries reside, these networks offer the framework needed to support ILL applications, and thereby eliminate the barriers to ILL communication which currently exist.

3.2.2 Electronic document delivery

Electronic document delivery refers to the use of electronic forms of documents and transfer media to support the transfer of documents between libraries, and between libraries and end-users. The technologies and international standards that have potential for use in representing electronic documents and in exchanging them over research networks will be discussed generally in this section. A more detailed examination of these technologies and standards, and their potential to form the "building blocks" of an integrated information system that could provide, not only immediate access to bibliographic information, but also online delivery of actual documents, is the subject of a separate UDT research project published as part of the *UDT Series on Data Communication Technologies and Standards for Libraries*.

Within a research network environment, electronic document delivery can be viewed as the complement to information retrieval; that is, as distributed information retrieval involves users at their workstations finding *references* to documents, electronic document delivery involves sending the actual *document* to their workstations. Currently, long delays in the delivery of documents is experienced because of a number of factors: the speed in forwarding the user's request to the lending or supplying institution, the speed of handling requests at the institution, the distance between the user and the supplier, and the method used in document transfer (Fjällbrant, 1988). Most documents today are supplied on paper and delivered by regular postal services. The British Library Document Supply Centre (BLDSC), for example, currently fills 60 per cent of its requests in this manner (Plassard, 1989) (although the majority of requests *sent* to

the BLDSC are now in electronic form (Dempsey, 1991a)). Reliance upon traditional methods of document delivery has created a bottleneck in the flow of information to the user, especially since access to bibliographic material has been substantially augmented through information retrieval. The quantum leap in bibliographic access to be provided by distributed information retrieval over research networks will make the disparity between information access and document supply even more severe. The balance can potentially be redressed by using research networks to send documents in electronic form directly to the user's workstation where it can then be read on-screen or printed off-line.

The technologies, application protocols, and standards available to support electronic document delivery over research networks include those for converting paper documents into electronic form, those for representing documents in electronic format, and those for actually delivering documents over networks. This section will discuss the technologies, protocols, and standards that fit into these three domains.

Document conversion

A major problem in shifting from delivery of paper documents to electronic documents is the extensive body of print materials that libraries have collected that must be converted into electronic form. Re-keying the materials, of course, is out of the question. Scanning and optical character recognition (OCR) technologies, discussed below, which automatically put materials into electronic format, can help speed the conversion process. Despite the use of scanning and OCR, the enormous volume of documents to be converted will make document conversion an expensive, time-consuming, and labour intensive undertaking, far surpassing retrospective conversion projects of the 1970's and 1980's.

Scanning Scanning is one of the techniques used to transform paper documents into electronic format. It produces representations of documents through raster-scanning that are simple bit-mapped images of pictures and/or text (Cawkell, 1987). Raster-scanning digitizes images on a page by converting its light and dark areas into 0s and 1s a strip at a time. The resulting binary representation is called a bit-map. The scanned picture is reconstructed from this information each time it is needed. A drawback of this form of storage is that the content of the electronic image cannot be edited or manipulated. For this reason, standard information retrieval techniques, such as keyword searching, cannot be used to retrieve relevant documents; bit-mapped images, unless processed further, are beyond the capabilities of these systems. Another limitation is that bit-mapped displays, even with data compression, require large amounts of storage space. It is the most useful technology, on the other hand, for *en masse* electronic conversion of paper documents. Pages of materials can be simply scanned, checked for quality, compressed by the system and stored. Probably the most familiar use of scanning technology to librarians are facsimile machines, or FAX, because of their use in interlibrary loan and current document delivery services. FAX machines digitize, compress, and transmit bit-mapped images over ordinary telephone wires that are re-constructed and printed at the receiving end (Cawkell, 1987).

OCR OCR, or optical character recognition, is a method by which the textual content of scanned images are converted into character code thereby making the text available for manipulation. In the most common method for character recognition, called *template matching* or *matrix matching*, the bit-mapped image goes through a process that compares the bit-map of a target character with a set of stored bit-maps; if the reader finds the target font represented among its repertoire, then the target font is recognized, otherwise, it is not (Cahan, 1989). The second method, called *pattern recognition* or *feature extraction*, is a more powerful method that analyzes the topological features of characters. It uses a number of algorithms or rules to go through a recognition process to identify the character. For example, an "O" contains 0 vertical lines, 0 horizontal lines, 1 loop, and 0 line endings, while a "D" contains 1 vertical lines, 0 horizontal lines, 0 diagonal lines, 1 loop, 0 line endings (Schein, 1989). A major advantage of this method is that the reader can be taught to recognize font types that it initially fails to recognize. The output of both processes is character code, such as ASCII.

The limitation of OCR technology is that the rate of accurate recognition must exceed 97 percent to be more efficient than manual re-keying (Walker, 1988). All errors produced in the OCR process must be corrected by hand. Factors that reduce recognition accuracy are poor paper or typeface quality; OCR systems require clean, crisp, black text on a white background to be most effective. Another limitation is that OCR technology also has problems with features of documents such as multiple columns, equations, and diagrams.

Yet, there are distinct benefits in using this technique over strict imaging. First, the amount of storage required for each page is reduced significantly since only ASCII code is stored. Second, text-based tasks, such as information retrieval and automatic indexing, can be used on this type of electronic material. Typically, however, projects that have begun mass digitization of print materials use a combination of bit-mapped images and OCR generated text to provide access to the material. One such project is the National Agricultural Text Digitizing Project (Andre, Eaton, and Zidar, 1988). In this project, the documents were stored optically in bit-mapped form, while indexes for the images were created using OCR generated-text with indexing software. Pointers linked the index entry with the appropriate page images. This combination of text and images provides both information retrieval capabilities and the advantage of viewing an "original" picture of a document.

Document representation

There is an increasing trend to represent documents in electronic format throughout all stages of their development, from initial creation on word processors, through editing, to final typesetting or on-screen display. This trend will supply ready-made electronic documents for future distribution. While all electronic documents are ultimately binary code, at a higher level, they can take a variety of forms. Major document representations are outlined briefly below.

SGML A step up from ASCII representations are those that have elements of the logical structure of a document, such as headings, subheadings, and paragraphs, identified within the body of the text. Standard Generalized Markup Language (SGML) is a device-independent, process-independent ISO standard for adding "tags" identifying these elements (Wu, 1989).

Other types of content elements, or *entities*, include headings, chapters, paragraphs, and appendices. SGML also allows for general rules that describe specific types of documents. These rules define the entities allowed in a given document type as well as the relationship among the entities. For example, a journal article can be defined as having a title, an author name, an abstract, an indeterminate number of headings and paragraphs in the body, followed by references. Tagging documents with SGML simplifies both information retrieval and automatic indexing of full text as the significant entities used in these processes are explicitly identified. Note, however, that SGML does *not* describe the actual *layout* of a document.

SPDL Almost the opposite of SGML, Page Description Languages (PDL), such as PostScript, define the layout of an electronic document but not the content; they specify the exact manner in which a page should appear when printed. They tell the printer how to size and position minute elements, such as lines and curves, that will form the fonts and graphic elements on a page. The ISO standard for PDLs is SPDL, or Standard Page Description Language, an attempt to create a PDL that is not exclusively tied to a specific text generation package or printer (Robinson and Strasen, 1989). Thus the same PDL description could be sent to a dot matrix printer or to 2400 line-per-inch typesetting machines, and each would be able to produce a page of output using its best resolution.

Note that documents defined with a PDL are in a final, non-processable form; that is, they cannot be readily manipulated or modified as other electronic representations. In this way, it may be a representation that will be useful in cases where the recipient has no need to modify the document. A PDL description, on the other hand, may not be suitable as a *general* transfer mechanism because the description describes only page layout, and not screen layout. If the document is required to be output on a device such as a terminal, which has a different page size, there will be compatibility problems. Similarly, there will be problems in exchanging documents between Europe which uses metric standard page sizes, and North America, which uses Imperial standard pages.

ODA Office Document Architecture (ODA) is a method of specifying both the content and layout of a document (Dawson and Neilson, 1990). The specific aim of ODA is to facilitate the interchange and remote access to documents (Radlett, 1988). The standard is used to define complex, structured documents that can simultaneously contain text and graphics. It may also contain audio and video content, although work on these elements is not far advanced. Essentially, documents produced using ODA can come in three forms: 1) the *processable form*, or revisable form, that permits the recipient to further process both the content and format of the document, 2) the *formatted processable form* that has its format fixed but allows the recipient to modify the content and, 3) the *formatted form*, or final form, unmodifiable by the recipient and ready for printing on-screen or off. ODA uses SGML to define the Office Document Language (ODL), allowing an ODA structured document to be represented in accordance with SGML.

ODA representations are the most sophisticated of those outlined above and have two distinct advantages over other representations. One, all aspects of a document needed for most tasks, such as information retrieval, automatic indexing, printing on-screen or on paper, are contained

in a single document. Two, any of the three document forms outlined above may be transmitted depending upon the task for which it is needed. For example, the processable form may be exchanged between authors for further modification and the final, formatted form may be distributed to end-users.

Document transmission

Because research networks support the exchange of data between computers, they have the potential to be used as the conduit for electronic documents. The same characteristics that make them suitable for information retrieval, that is, their interconnection of computers that store information and users' ability to access them from wherever they work, make them suitable for delivering actual documents. There are network services and standards currently available, or under development, that can be used to support the delivery of electronic documents. The functionally simpler of two major transfer methods is remote file transfer. More promising, however, is the use of message handling systems to send all forms of electronic documents directly to users.

File transfer Recall that file transfer is the network service used for transferring files between computers. Within TCP/IP and OSI realms, the protocols for file transfer are FTP and FTAM, respectively. This simple method of transferring documents from computer to computer involves connecting to a remote host and requesting that a particular file be sent or received. A limitation of this method with regard to document delivery is that a direct connection between machines is necessary for the file transfer process; documents cannot be simply sent to the user to await acceptance at a later unspecified time. This is particularly a problem with the transfer of image data which can involve very large files and transfer times.

Message handling systems and document delivery Message handling systems based on the X.400 application protocol manage the delivery of mail by providing an addressable envelope into which users, machine or human, can insert messages that are then forwarded by the system. Of special significance for electronic document delivery is the fact that these systems allow a wide variety of message types to be enclosed within the envelope (Swain and Tallim, 1990). The two forms perhaps most germane to this discussion are bit-mapped facsimile images and documents structured with ODA. The ability to accommodate these types of document representations makes X.400-based message handling systems particularly useful for building applications for document delivery. Normally, fax images are exchanged through the use of dedicated fax machines permanently connected to their own telephone line. In message handling systems, the fax image can be contained within the envelope thereby allowing its transmission over an X.400 network instead of a telephone line. The ability to accommodate ODA documents within the envelope allows such documents to be sent over a network to any recipient, human or machine, for further processing or printing.

Employment

The standards and technologies outlined above form part of the building blocks suitable to create integrated document delivery systems to deliver documents directly to users. Large databases created from the existing paper corpus, as well as databases of future documents structured with

SPDL, SGML, or ODA, can serve as sources of electronic documents. Message handling systems, such as those based on the X.400 standard, have the potential to be used to both send the request for a document to libraries or document supply centres automatically, as well to deliver the actual document, in whatever form, back to the requestor. Other functions of an integrated system that MHS may support are payment requests, delivery notices, and error messages (Tuck, 1989). The use of MHS for these functions would reduce the need for costly re-keying. In this way, the bottleneck that currently exists in manual document delivery systems could be eliminated. Staff would no longer have to find the paper original, make a copy, and send it by post. Instead, the request for a document could be automatically processed by calling up a previously stored electronic representation of a document and sending it to the user by reply email.

The proper function of integrated document delivery systems may be document *distribution* as opposed to document *interchange*. Document interchange involves the exchange of electronic documents in a form that can be revised and manipulated, such as between co-authors, while document distribution involves electronic documents that are in some final, non-revisable or unmodifiable form (Robinson and Strasen, 1989). Information sent to an end-user via MHS should no doubt be in this latter form to prevent unwanted modifications. While exchanging processable forms of documents between networked authors is a valuable use of these technologies, it may be that for information provision, only documents in bit-mapped, SPDL, or ODA formatted forms will be appropriate for distribution to end-users. However, developments of products that can manipulate even so-called "final forms", such as those that can perform OCR on a FAX file, may make any particular final form of a document an ephemeral entity.

A major stumbling block for electronic document delivery in any form are the issues of copyright and royalty payments. Scanning existing copyrighted material, for example, violates the copyright. The whole concept of delivering documents electronically is based on making a copy of a work every time it is needed. In a research network environment, where literally millions of copies of a work may be sent to millions of users, developing strategies for fair compensation becomes crucial. Significant work must be done to solve these problems before any technology, no matter how sophisticated, can be used to store and deliver documents via research networks. See section 5.2.2 for more information concerning copyright.

3.3 Applications for technical services

The third domain of library work with applications to research networks is technical services. Technical services traditionally encompasses all of the staff-oriented, behind-the-scenes tasks in which libraries engage such as collection development, acquisitions, serials control, cataloging, shelving, and preservation. These activities ultimately make bibliographic information provision

and document supply possible; collecting, cataloguing, and classifying materials, creating and maintaining machine-readable databases, all support making information accessible to the user. The expensive and time-consuming nature of these tasks has made them prime candidates for advances in computer automation from the earliest punched card equipment, to integrated library systems, to the current generation of research networks (Arms, 1989a).

A discussion of technical services and research networks would be incomplete without reference to the many library networks that currently exist to support these functions. Libraries recognized the utility of some form of computer network to support technical service and resource sharing activities early in the history of library automation. These networks were developed primarily to support shared cataloguing, although other services were eventually offered, such as acquisitions, serials control, collection development and ILL. Major library networks include OCLC (Online Computer Library Center), RLIN (Research Libraries Information Network), and WLN (Western Library Network) in the United States, Dobis and UTLAS in Canada, and BLCMP and SWALCAP in the United Kingdom. These networks, however, are mainly closed networks; that is, they are comprised of terminal-to-host links, dedicated terminals, centralized hosts, proprietary protocols, and idiosyncratic user interfaces. They have been traditionally a series of distinct networks that are unable to communicate and directly exchange data, either with their respective members or among themselves (Avram, 1988b).

There have been efforts to link U.S. systems into an open architecture through the Linked Systems Project (LSP). The LSP chose the OSI model to attempt to connect the disparate library networks into one national peer-to-peer library network, and to date has succeeded in establishing some links for specific tasks (Avram, 1988b). While this project may eventually satisfy libraries' needs for a national library network for support of technical services, a number of points suggest that research networks are a viable alternative to library networks. They are that:

- research networks *already* offer a general peer-to-peer communications infrastructure that can be utilized in the exchange of library processing data. This infrastructure, characterized by large interconnected systems such as the *Internet*, JANET and NORDUnet, has been developing over the last 20 years and is evolving into a single, completely interconnected world-wide network (Quarterman, 1990)

- the institutions where research network hosts reside also tend to have libraries that use library networks. At least for academic libraries, this situation threatens to split library activities between two parallel networks. Instead, applications should be developed to support technical services over the existing research network infrastructure along with applications for information provision and document supply. *One* logical network should serve the needs of the entire library community, including both users and staff (Avram, 1988a)

- the LSP is based on pay-by-the-packet public packet-switched networks (*Telenet*) rather than its own physical network. This may eventually create funding problems, especially as file size and network activity increases (Lynch, 1988)

- to realize the full potential of application protocols for technical services, reliable, affordable, high-capacity telecommunications are required (Buckland and Lynch, 1988). Research networks are now a mature technology that are increasing in both speed and bandwidth.

Libraries outside the U.S. have begun to use the communications resources provided by research networks to support technical services. In Britain, Europe, and Japan, for example, the activities that were previously carried out on indigenous library networks are supported by their respective research networks, or are planned for integration in the near future (Stone, 1990b; Buitelaar, 1990; Naito, 1989).

The application protocols developed within the LSP, such as those for the retrieval and transfer of cataloguing records, are examples of the types of application protocols that can potentially support technical services over research networks. Although these protocols are OSI-based, the planned migration of TCP/IP networks to OSI and intermediate strategies such as OSI-gateways or OSI running over TCP/IP, make them applicable to these networks. Other protocols and standards, particularly X.400 and EDIFACT, have uses in technical service applications such as acquisitions. These applications and standards are discussed below.

3.3.1 Shared cataloguing

Foremost among the technical service applications of research networks is shared cataloging. Shared cataloguing refers to the cooperative exchange of records among libraries thereby reducing redundant original cataloguing and related costs; a library does not need to create a cataloguing record from an original work, but can copy records that have been created at another institution (Avram, 1988b). Historically, the sharing of cataloguing records began with the distribution of catalogue cards by the Library of Congress. During the 1960's and 1970's, these records were converted to machine-readable form based on the MARC format and stored in centralized databases. Bibliographic utilities, such as RLG, OCLC, and WLN, developed networks to provide access to these databases allowing member libraries to contribute new records and to draw others from this central store. The dominant media for record transfer, however, remained a combination of magnetic tape and terminal-to-host telecommunications. In a effort to increase the efficiency of record transfer, the LSP was initiated to link library networks in a unified peer-to-peer network that would support the direct machine-to-machine exchange of cataloguing records and other library processing data (Avram, 1988b).

Out of this project has come a number of application protocols that can be used to support applications in shared cataloguing, specifically, the machine-to-machine transfer of cataloguing and authority records. As mentioned above, they have applicability for use on research networks, especially as research networks migrate toward OSI protocols. They will be discussed here, however, within the context of their use in the Linked Systems Project.

LSP application protocols

Information Retrieval This is the original protocol for the exchange of bibliographic information upon which NISO Z39.50 and ISO SR are based. These later versions were extended to support information retrieval applications, in general, and thus has application in the area of information provision. Within the context of the LSP, however, the Information Retrieval protocol supports the exchange of cataloguing and authority records.

To review, this protocol enables searches to be carried out between disparate, remote systems using the database of one system and the search interface of another; the user interacts with the local interface with same search syntax and result set display, while in the background the system accesses the databases of heterogeneous remote systems. The operation of the *Information Retrieval* protocol between the Library of Congress (LC) and RLIN, for example, would proceed as follows (from McCallum, 1988:13). During a search for a name authority, a cataloguer at LC would enter a normal file specifier at the end of a search statement, such as *file=na*, modified slightly to specify a remote file, *file=rlgauth*. The local system would then translate this statement into an intermediate syntax, which would then be sent to the RLIN system and converted into RLIN syntax. Using its local syntax the remote system would proceed with the search.

The information resulting from the search can be returned to the local system in one of three ways: 1) if a set is less than a size specified by the LC system, all records are returned, 2) if a set is less than a larger specified size, both a total count and sample records are returned, and 3) if the set is larger still, only the total count is returned. Providing total counts and sample records in cases where the set is relatively large allows the cataloguer to limit the search further. The final display of name authorities is, of course, in a format native to the local system.

Record Transfer The record transfer application protocol is a file transfer protocol used to exchange groups of records of various sizes between systems (McCallum, 1988). It allows an *initiator* system to request the transfer of records from a sequential file residing on a *responder* system. It is roughly similar to FTAM, except that FTAM allows the initiator and responder to either send or receive data, while in Record Transfer the initiator is always the receiver and the responder is always the sender (Denenberg, 1988).

LSP implementations

These application protocols have been used in the LSP Authorities Implementation designed to support the exchange of authority records between selected libraries (Avram, 1988b). This implementation has provided the National Coordinated Cataloguing Operations (NACO) program with more efficient method to transfer data. The NACO was created in 1977 by the LC to create a national database of authority records in an effort to share resources and provide consistent use of preferred forms. Prior to the Authorities Implementation, participating libraries submitted authority records to LC on worksheets mailed to the library and through terminal-to-host connections. Currently, however, cataloguers can submit records to their utility for subsequent machine-to-machine transfer to LC's authority file. These data can then be distributed among all participating utilities. RLIN and LC have been exchanging records in this fashion for more than

four years, while OCLC has been linked for three years. This method of transfer has eliminated time-consuming tape loading of authority records by the utilities (Avram 1989b).

Work is currently progressing to make Information Retrieval and Record Transfer support the exchange of cataloguing records. This work will take place within the National Coordinated Cataloging Program (NCCP), a program similar to NACO, in which selected libraries will participate in creating a national database of cataloguing records. Procedures for the transfer of cataloguing records will correspond to those used in authority records transfer with uploading to the LC central database and subsequent distribution.

Employment

The use of these protocols has the potential to provide individual libraries with the ability to obtain cataloguing records from a variety of sources. Akin to distributed information retrieval, where an individual user can search a variety of databases on a network as if it were a single database, individual libraries could potentially use applications based on these protocols to search and retrieve cataloguing records from other libraries. There would, of course, be a need for widespread adoption of appropriate application software before this could be accomplished. There are also copyright restrictions regarding the sharing of cataloguing records that must be considered. OCLC's copyright claim on its union catalogue, for example, restricts libraries using these records from sharing them with libraries that do not have a licensing agreement with OCLC (Lowry, 1987). Much work would have to be done in this area before such applications could be used effectively.

3.3.2 Acquisitions

The second potential technical services application of research networks lies in the area of acquisitions. Acquisitions is the process by which libraries obtain materials for their collections, primarily through purchases from jobbers, book sellers, and publishers. While much of the work involved in automating acquisitions has focused on managing various "housekeeping" tasks, such as order monitoring, receiving, and fund accounting (Chamberlain, 1985), there have also been efforts to develop direct telecommunications channels with suppliers. These channels facilitate the exchange of business data such as purchase orders, invoices, and receipts, that are generated during the acquisitions process. Currently, there are various methods by which these data are exchanged. For example, many libraries order materials directly from book suppliers through private electronic mail systems (e.g. Envoy in Canada) or by fax. Some larger vendors (e.g. Brodart), support full electronic ordering systems that handle transactions in electronic format (Cline, 1990). Bibliographic utilities have also provided connections that allow direct electronic transmission of orders to participating vendors. Further, some integrated library systems have subsystems that provide electronic links vendors (Walton and Bridge, 1987).

Despite the greater efficiencies that the electronic exchange of business data between libraries and vendors has provided over manual methods, limitations still remain. These limitations arise from the heterogeneity of message formats and of the telecommunications channels that carry them. In this way, the exchange of acquisitions data closely parallels the exchange of ILL

messages; each method of communication, whether a vendor's own proprietary ordering system or a bibliographic utility's subsystem, has a different message format associated with it. This situation results in barriers to efficient business communication similar to those that affect ILL communication. These barriers are that:

- separate telecommunications channels prohibit the exchange of messages between different electronic ordering systems
- conflicting message formats prevent automatic processing by dissimilar systems.

These barriers require libraries to maintain connections to a variety of systems to take advantage of the many sources of material available for acquisitions.

In order to overcome the barriers to acquisitions communications, a standardized message format for acquisitions data, and one communications channel for the exchange of messages will be required. Standards developed outside the library community for the exchange of business data have the greatest potential for accomplishing this task. Specifically, these are the standards that fall under the general heading of Electronic Data Interchange, or EDI. An explanation of EDI and its potential for use in acquisitions is outlined below.

Electronic Data Interchange (EDI) EDI is defined as a "series of standardized message formats, primarily oriented toward business transactions, used for inter-system communications" (Swain, 1990:204). In effect, EDI allows the computer-to-computer exchange of business documents between two companies in a standard electronic format. EDI was developed to eliminate the use of paper documents and to facilitate fast, efficient, and accurate information flow. Initially, the transmission of electronic business data remained closed within specific sectors, such as the grocery, banking, and financial sectors, with little cross-sector interaction. There are, for example, 15 different spector-specific standards for EDI in the U.S. In general, EDI formats consist of three components:

- a data dictionary which defines the content of data elements used in messages
- a set of message standards which define the format and data content of messages for specific business transactions
- a message syntax which defines the structure of the messages and additional information needed to exchange data.

(Swain, 1990:211).

In the early 1980's, work began to develop EDI standards that could be exchanged across industry sectors and national boundaries. Within the U.S, the American National Standards Institute (ANSI) created a standard for cross-industry EDI known as *ANSI X12 Data Interchange Standards.* The European EDI equivalent is the called the *United Nations/Trade Data Interchange*, or UN/TDI. Efforts to create an international EDI standard began in 1986,

resulting in EDIFACT (Electronic Data Interchange for Administration, Commerce and Transport) which combined elements from both the ANSI X12 and UN/TDI standards. Currently ANSI X12 and EDIFACT are the dominant EDI standards, with X12 used primarily in the U.S. and EDIFACT in Europe (UDT Newsletter, 1990).

EDI and X.400 EDI and X.400 have complementary attributes that make their conjunction particularly useful for the exchange of business messages. On one hand, EDI standards define the content and syntax of business messages, and not the underlying communications technology used for exchanging messages between systems. On the other hand, X.400 is a "generic" message handling technology that supports the exchange of multiple message types regardless of content (Swain, 1990). In this way, EDI messages, like ILL messages, can be enclosed within the X.400 MHS "envelope" and forwarded to their destination. The major benefit of the use of X.400 is that it provides the ability to exchange EDI messages between systems within an open systems environment thus overcoming the boundaries created by numerous proprietary systems. Furthermore, the store-and-forward approach used by X.400 is ideal for the asynchronous nature of business communications which, moreover, often take place across different time zones.

The CCITT (International Telegraph and Telephone Consultative Committee) has recently approved changes in the X.400 standard to allow the inclusion of EDI messages as a body part. The so-called EDI-over-X.400 service allows greater flexibility in defining the additional service required by EDI applications, while using existing X.400 services as much as possible. These changes are contained within CCITT Draft Recommendation X.435 which is currently undergoing the CCITT balloting process (CCITT Draft Recommendation X.435, November, 1990).

EDI and Libraries

While most of the activity surrounding EDI and X.400 is taking place outside the library community, the standardization of messages and transport media has great potential for application to the acquisition process. Because the purchase of library materials involves the exchange of pricing information, order forms, invoices, and receipts, X.435 will be as applicable to acquisitions transactions as to non-library business transactions. Acquisitions departments will be able to profit from the same benefits that EDI brings to the business world in general. These benefits include:

- increased efficiencies in both internal operations and external communications
- reduced paper work
- fewer transcriptions errors
- reduced overheads
- improved information flow and turn-around times.

The use of research networks to support EDI has the potential to allow libraries to engage in acquisition transactions with multiple vendors through one convenient communications channel, eliminating the need to maintain separate channels to each vendor. The full potential of research network use for acquisitions processes, however, is limited by a number of factors, at least in the short term. These factors are: 1) research network connections are currently available only to libraries located at university campuses and research institutions, thus excluding many public and special libraries from an integrated system, 2) the connection of jobbers, book sellers, and publishers to research networks may violate "appropriate use" guidelines that prohibit commercial use of the networks (see 5.2.3). Definitions of appropriate use, however, are changing rapidly. Recent efforts to commercialize the *Internet* is a case in point (see Jackson, 1989). This is a major issue that would have to be resolved before research networks could be used to support EDI in acquisitions processing.

3.4 Other library uses of research networks

Other uses of networking, while of less immediate relevance to the operational concerns of libraries, are precisely those areas in which research networks are most highly developed. They offer considerable potential to librarians for professional contact and information exchange as well as to provide additional services to their clientele.

3.4.1 Bulletin boards

Electronic bulletin boards act much like regular bulletin boards which provide a public space to post notices. Libraries can and do use them to provide patrons with local information about the library such as opening times, borrowing policies, bibliographic instruction schedules, library events, and recent accessions. Bulletin boards, when they are offered, are often part of the online catalogue. Network wide, such information would be useful to assist individuals in planning a trip to a far away library and for providing information about the network itself. JANET network users, for example, can access a bulletin board news service that provides information about JANET such as a list of JANET-accessible and addresses, brief catalogue content descriptions and login procedures, a bibliography on JANET use, and a directory of library and staff email addresses.

3.4.2 Electronic conferences

Electronic conferences, or discussion lists, are an application of electronic mail that support on-going discussions on specified topics among widely dispersed colleagues and, as such, are powerful tools for professional communication. Individuals within a particular interest group can make comments, request information or help, announce new procedures or techniques, and generally keep up to date by participating in a conference. The conferences themselves are managed by software such as LISTSERVers on BITNET. A LISTSERV maintains a list of email

addresses of conference participants. It essentially acts as a distribution node for all messages sent to it; a message sent once to a LISTSERV is automatically sent to all addresses on a particular mailing list. In this way, continuous discussions are generated among any number of colleagues.

There are currently hundreds of discussion lists active on BITNET on a host of topics from vegetarianism to nuclear physics. Libraries have also begun to develop conferences on library-related topics. The list below represents some examples of library-oriented discussion lists available on BITNET. Anyone with access to a BITNET/INTERNET account may subscribe to them. A few examples include:

Discussion list	Topic
ARLIS-L@UKCC	Art Libraries Association of North America
ATLAS-L@TCUBVM	Discussion list for user's of Data Research's ATLAS system
AUTOCAT@UVMVM	Library Cataloguing and Authorities
BI-L@BINGVMB	Bibliographic Instruction
BRS-L@USCVM	BRS/Search Users
BUSLIB-L@IDBSU	Deals with all issues related to the collection, storage, and dissemination of business information in a library setting
CWIS-L@WUVMD	Campus-wide information systems
ELLASBIB@GREARN	Library automation systems in Greece
EXLIBRIS@RUTGERS	Rare books, manuscripts, and special collections
GOVDOC-L@PSUVM	Government Documents
INNOPAC@MAINE	Innovative Interfaces Users
IR-L@UCCVMA	Information Retrieval Research
LAW-LIB@UCDAVIS	Law librarian discussion list
LIBPLN-L@QUCDN	Library Planning
LIBREF-L@KENTVM	Discussion of Library Reference Issues
MLA-L@IUBVM	Music Library Association
NOTIS-L@TCSVM	NOTIS Users
NOTISACQ@ULKYVM	Notis Acquisitions
NOTMUS-L@UBVM	Notis Music Librarians
NOTRBCAT@INDYCMS	Notis Rare Book Catalogers
PACS-L@UHUPVM1	Public-Access Computer Systems Forum
SERIALST@UVMVM	Serials Users Discussion Group
SPILIB-L@SUVM	SPIRES Users

(Source: collected from PACS-L correspondence, 1991)

4. LIBRARIES AND RESEARCH NETWORKS: PROJECTS

Although the development of research networks dates from the late 1960's, the explosion in the growth and interconnection of research networks did not begin until the mid-1980's. At about the same time, libraries began to connect their systems to research networks in order to supply their services to this expanding body of network users. The British academic network, JANET, for example, began providing access to library catalogues as early as 1985, barely a year after the network's inauguration. There are currently major projects linking libraries and research networks underway in the United States, the United Kingdom, the Netherlands, Sweden, Finland, Australia, and Japan.

An examination of these connectivity projects reveals a number of overall patterns. Recall from the previous chapter that library uses of research networks can be divided roughly into three categories, bibliographic information provision, document supply, and technical services. The first pattern that emerges is related to these divisions; that is, most of the work involving research networks thus far has centred around the first category, providing network users with access to information resources. Major networks, such as JANET and the *Internet*, for example, probably have connected a combined total of over 260 library online catalogues by the spring of 1991. The resources, however, remain largely limited to databases of this nature. Less progress has been made in the other two categories, the support of document supply and technical services over research networks. Major bibliographic utilities in the U.S., for example, seem reluctant to use research networks to supply the technical services currently available to libraries over their own private networks. The general trend seems to be that in areas where extensive bibliographic and library networks have developed to support technical services, these services have not been integrated into the infrastructures of research networks. In other parts of the world, such as Britain, Europe, and Japan, planners have started to include these functions in their development projects.

A second major pattern is that standard network services and other applications that would enhance library use of research networks in all library-related work are either not currently used by libraries, or are under development. These applications include standard file transfer, as well as more advanced services such as directories, distributed information retrieval, and the use of standardized interlibrary loan protocols. For bibliographic information provision, lack of development in this area causes problems in finding what information resources are available, searching them with consistent interfaces, and directly downloading search results. For document supply, this lack results in no standardized way to order ILL requests or deliver electronic copies of documents. For technical services, the problem is that there is currently no way on most networks to send cataloguing data directly to the cataloguer's workstation.

Clearly, the seamless access to information resources that has been envisioned by many authors, and the utilization of research networks for all library functions, is still some distance away. Much work is needed in integrating all library applications into the infrastructures of research networks, and in exploiting the enhanced services that research networks can provide. Still, the sheer number of library and research network projects underway, and the pace with which they are advancing, is promising. Unfortunately, this rapid pace of change also makes it difficult to obtain detailed up-to-date information about many of the projects. Nonetheless, this chapter briefly outlines the major projects that are in progress, and discusses the above patterns in relation to these international developments.

4.1 Library and research network projects in North America

The research network to which libraries in the U.S. are connecting is the *Internet*, a TCP/IP-based internetwork that uses the DARPA and NSFNET networks as backbones to connect a large number of smaller component networks. Although most of the mid-level networks connected to the *Internet* are found in the United States, the network also extends beyond North America to Europe, Japan, Australasia, and Mexico. Because of the *Internet*'s international scope, the systems connected are not only American systems, but also include various Mexican, European, and Australian library systems.

4.1.1 Library services on the *Internet*

The overall picture in the U.S. seems to be one of a collection of disparate library systems connected to the *Internet* in an *ad hoc* fashion. Libraries provide simple login access to their systems, but still rely on established channels for technical services. Major American bibliographic utilities do not at present use the *Internet* to support shared cataloguing and other technical services; these services are only available through their private networks. In the end, there is no real integration of library functions, both technical and end-user-oriented, into the *Internet* infrastructure, nor is there widespread *library* use of standard network services. The full potential of the research network remains to be developed.

Nevertheless, any networker with an account on the *Internet* can currently access over 200 different library catalogues of major American universities as well as a number of collections abroad. A researcher sitting at her desk at SUNY Buffalo, for example, can view holdings of a university library in Germany, or in California, or a neighbouring state, or in many other locations. The number of catalogues on the *Internet* has grown from only a "few" in 1988, to 30 in 1989, to 40 in early 1990, and then to its current level of over 200 as of Spring of 1991 (Lynch, 1991). Although this represents a vast increase over the last three years in the information resources available to researchers from their desks, the connections remain fairly crude. It is up to the user to discover the locations of the catalogues, to connect to the remote systems, and to learn the appropriate commands.

Detailed descriptions of the catalogue contents, login procedures, and search methods of the library systems available over the *Internet* are beyond the scope of this chapter. For more detailed information concerning connecting to *Internet*-accessible catalogues and their *Internet* addresses, see one of the following references (Barron, 1990; St. George and Larsen, 1990; NSF Network Service Center, 1990). For additional information concerning library/research network connectivity projects in North America, in general, please see Dempsey's (1991b) recent review, as well as an article by Lynch and Preston, (1990).

It is useful, however, to give a broad overview of the types of library services that are available on the network, how they are accessed, and the problems encountered when attempting to use them. In this way, one can appreciate the work that has been done so far in using research networks for library purposes, as well as gain an understanding of the work that still needs to be done.

Access

Directories There is, at present, no formal directory structure to guide the user to a particular catalogue or database on the *Internet*. Awareness of the networked catalogues, unfortunately, tends to be mainly by (electronic) word of mouth. There are several documents available that act as resources for accessing *Internet* library catalogues. For example, one is available by sending the message GET INTERNET LIBRARIES to LISTSERV@UNMVM.BITNET. This will retrieve a large document by St. George and Larsen (1990) that provides most information a user would need to connect and search remote catalogues.

While documents such as those mentioned above are useful interim solutions, a proper online directory structure is required to fully utilize the burgeoning resources available on the *Internet*. NYSERNet, Inc. is currently involved in a pilot project to develop an OSI based "White Pages" service. This is, however, is a voluntary, informal project designed to determine the type of directory service that will be needed by users, and to discover the limitations of an X.500-based directory (Rose and Schoffstall, 1990). It is also not equivalent to a "Yellow Pages" directory, the type of directory that would be needed to located resources according to subject.

Connection Most of the hosts of the library systems on the *Internet* are accessible via TELNET, the TCP/IP protocol for remote login. Through TELNET, users can directly connect their own terminals to the remote library system. A few other machines require TN3270, the version of TELNET for IBM 3270 terminals. Connecting to the remote system entails dialing the system's address with the TELNET command. This address can take two forms, a *domain name* address, and its numeric counterpart, the *IP* address. An example of a domain name address is *nerm.nerdc.ufl.edu*, the address of the library system at Florida State. This address expressed as an IP address is 128.173.5.4. While the domain name is most common, a few hosts do require the second address form.

Login procedures Finding the proper address and connecting to the host is usually only the first step in the process of accessing a catalogue. Login procedures of varying complexities are often required following the initial connections. These can range from as simple as typing the

name of the library system to the initial prompt, to as complex as a sequence of commands, escapes, clear screens, and carriage returns.

An example of how complicated logon procedures can become is that required to access Florida State University Library System's LUIS. A number of steps are needed to get into library system. They are:

- Type TN3270 nervm.nerdc.ufl.edu to connect to the host
- Press <CR>
- Type DIAL VTAM to active VTAM service
- Type NERLUIS on the screen which displays the message NERDC VTAM IS ACTIVE
- Enter FCL as the GROUP-ID on the security module
- Enter AAYA as the OPERATOR
- Enter AAYA again as the PASSWORD
- Press <CR>
- Press KEYPAD ENTER when the SIGN-ON COMPLETE screen is displayed
- Type LUIS to select the LUIS system

Further, many systems do not allow users to logoff once they have completed their tasks. In these cases, users must abort the session using a command such as @R (Lynch, 1989b).

Some systems require user IDs and passwords, as well, although these tend to be generic forms that are freely given to interested users or published in manuals. While experienced networkers may accept detailed login procedures as routine, a library user accustomed to walking up to a dedicated terminal in the library may find a long login sequence somewhat bewildering at first.

Terminal emulations The most commonly supported terminal is VT100. Other systems support only IBM 3270, requiring the use of the special version of TELNET, TN3270. Still other systems provide a list of standard terminals from which the users can select the most appropriate terminal type. For example, the University of Oregon's library system, JANUS, provides a choice of VT100, Wyse, TVI925, and emulator terminals. This wide variety of terminals can create many problems resulting from errors in communication between a host and a terminal that the host does not support. For a full discussion of the problems that can arise in the multi-terminal environment of the network, please see Lynch (1989b).

Search systems and interfaces The library systems connected to the *Internet* represent the range available from most major commercial vendors. These include NOTIS, Carlyle, IBM DOBIS/LEUVEN, and INNOPAC. This array of interfaces can create some difficulty for users if they are unfamiliar with the particular system used at a given university. In general, however, help screens are provided which can help steer users through an unfamiliar interface. The access points available on the various systems are, for the most part, the familiar points of author, title, and subject heading.

Information resources

Online catalogues The information resource most common on the *Internet*, as previously stated, is the online catalogue. Typically, these catalogues contain the records of holdings of a single university library or the combined catalogue of a large association of libraries. For example, ILLINET Online, the library system of the University of Illinois/Urbana-Champaign, includes the holdings of University of Illinois, as well as 800 academic, public, and special libraries throughout the state.

The scope of the *Internet*-accessible online catalogues, in general, includes monographs, periodical titles, films, videos, maps, manuscripts, and software. The extent of coverage varies widely, however, from catalogue to catalogue. For the majority of catalogues, coverage begins in the mid 1970's when most automated systems were initiated. The material represented by a given online catalogue can range as low as 40 percent and as high as 100 percent (derived from St. George and Larsen, 1990). A problem that arises with remotely accessing such databases is that the searcher may not know if he/she is accessing a full collection or only the portion that appears in the online catalogue. Most libraries have stated, however, that retrospective conversion projects of older parts of collections are in progress.

Non-catalogue information resources A number of libraries offer other types of information resources in addition to their catalogues. These resources include local library and city news, campus events, inhouse bibliographic databases, information about new journal issues, map databases, and information about government documents. Other non-restricted databases are also available such as CARL's (Colorado Alliance of Research Libraries) *Choice Book Reviews*, a collection of book reviews for college libraries, and special collections such as the University of Texas Austin Library System's Latin American, Middle East, Asian, and South Asian collections.

Some of the most potentially useful information resources that many university library systems offer to their local users, however, are not available for public access. These are commercial bibliographic databases that have been licensed and tape-mounted locally such as MEDLINE, PsychInfo, ERIC, Magazine Index, ABI/INFORM, and Chemical Abstracts. The licence agreements that the universities have negotiated with the producers of these databases restrict their use to a single campus or university library system. While they are available on the *Internet*, they are not available for general public access (Lynch, 1991).

The University of California's MELVYL system, for example, has licensed MEDLINE for use by its community, but does not provide public access to it through the *Internet* (Lynch and Berger, 1990). Rather than restrict access to MELVYL in general, only access to the MEDLINE function has been restricted through the use of a password.

Another database available on the *Internet* is CARL's *Uncover* database, a journal article database taken from the contents pages taken from thousands of multi-disciplinary journals. It will soon contain 10,000 titles and descriptions of 600,000 articles. CARL makes the *Uncover* database available over the *Internet* to institutions on a flat rate basis.

Restricting access to databases raises issues of authorization and access control. *Authentication* involves determining the identity of a user, while *access control* involves providing or denying access a resource based on that identity (Lynch, 1989b). Authentication and access control form the foundation of the protection and management of information and system resources. These issues are covered more fully in Section 5.2.5.

Availability The online public access catalogues of the libraries on the *Internet* are available to anyone who has access to the *Internet*. The citation databases restricted by licensing agreements are not available to all users, as discussed above.

4.1.1.1 Bibliographic utilities on the *Internet*

Although bibliographic utilities in the U.S. have been operating their own private networks since the early 1970's, efforts have been made recently to establish links between these networks and the *Internet*. The services that bibliographic utilities currently make available on the *Internet*, however, are not technical services. While the original focus of the bibliographic utilities was to support library technical processes such as cataloguing, acquisitions, in addition to interlibrary loan, this focus is shifting to bibliographic information provision. Both OCLC and Research Libraries Information Network (RLIN), discussed below, now provide TELNET access to a variety of databases. The technical services that are the traditional domain of bibliographic utilities remain accessible only through the private networks.

OCLC

OCLC, or the Online Computer Library Center, has been a mainstay of technical services departments in libraries in North America and the world for much of its history. It began in 1967 as a shared database of cataloguing information of Ohio libraries and has now grown into a major utility maintaining a database of nineteen million bibliographic records. It provides a variety of services such as shared cataloguing, serials control, acquisitions, and interlibrary loan (McGill and Racine, 1990). It also operates a private X.25 network connecting 9500 terminals and workstations worldwide. In recognition of the need to supply services to the *Internet* community, OCLC has developed the EPIC service.

OCLC's EPIC service OCLC's shift of emphasis from support of technical processes to bibliographic information provision has prompted the development of the EPIC service. The EPIC service is the information retrieval interface for OCLC's Union Catalog as well as a number of other databases. Users can access the service through the *Internet* at the address *epic.prod.oclc.org* (IRLA, 1991). Unlike *Internet*-accessible library catalogues, however, access to the EPIC service involves a usage fee.

The databases that are available with the EPIC service include:

- *OCLC Online Union Catalog*, the existing OCLC central database of nineteen million unique records on monographs, serials, manuscripts, music scores, machine-readable databases, software, and sound recordings
- *Twain Series*, a planned database containing 150 books of critical reviews of American literature in full text and graphics
- ERIC, an educational database
- *Book Data*, contains descriptive text and tables of contents of books from English language publishers
- *OCLC Union Catalog Practice Database*
- ABI/INFORM, a business database.

EPIC configuration The current EPIC search interface is based on the NISO Common Command Language (CCL) standard for Interactive Information Retrieval (Z39.58) (Helfer, 1990). There is also a workstation interface currently under development based on the Z39.50 information retrieval protocol. Under the Z39.50 specification, the information retrieval application is divided between the workstation, allowing its computing resources to be used, and a remote host where the database resides. In this way, an application on one computer can search the database kept on another. These separated applications are called origin and target applications; origin applications represent the client (the initiator of the query) and target applications represent the database server. The Z39.50 would allow the OCLC's EPIC Reference Service to be a target application to be searched by a client application located in the remote searcher's workstation (McGill and Racine, 1990).

Technical services There are presently no plans to connect OCLC's technical service functions to the *Internet*. This is the result of OCLC policy, rather than any technical barriers.

RLIN

The Research Libraries Group (RLG), created initially in 1974 as a consortium consisting of Harvard, Yale, and Columbia universities, is currently a non-profit organization governed by 36 major U.S. universities and research institutions (Richards, 1990). RLG operates a private terminal-to-host network, the Research Libraries Information Network (RLIN), which supports technical service applications such as cataloguing, acquisitions, and interlibrary loan. Similar to OCLC, RLG is in the process of changing its role from strictly a tool for technical services to one as an information provider. RLG has also recognized the importance of academic and research networks, and began making its databases available for searching over the *Internet* in May, 1989. It does not, however, support its technical services over the network.

Internet access In 1988, RLIN replaced its leased-line network with a private packet-switching network based on X.25 protocols with nodes stretching across the continental U.S. The RLIN host is connected to the *Internet* through BARRNET, an regional network in the San Fransico Bay area. Telecommunications for access from outside the United States must be arranged by the user, who must contact their local public data network to arrange access to RLIN through SprintNet. *Internet* access is free of telecommunication charges in the U.S., but may incur charges for access originating outside this area (Chapman, 1991).

For search-only access to RLIN there is a one-time "start up" fee of $200. Thereafter, use is charged on a per-search basis at $0.77 per search. Purchasing a block of 5000 searches brings the per-search charge down to $0.71 (Chapman, 1991).

Because the *Internet* is TCP/IP-based, and will be for the foreseeable future, RLIN has plans to support workstations that run OSI applications over TCP/IP (Richards, 1990).

RLIN databases RLIN presently offers a number of databases for end-user searching over the *Internet*. The software used to manage the databases is SPIRES, originally developed at Stanford University. Example databases include:

- *RLG Union Catalog*, a database containing 13 million unique records reflecting the catalogued holdings of its member institutions. The database is divided into eight sections pertaining to books, serials, visual materials, maps, musical scores, sound recordings manuscripts, letters, memorabilia, machine-readable files, and software

- *Avery Index to Architectural Periodicals*, a database indexing over 1000 journals in the field of architecture

- SCIPIO, an art sales database containing 88,000 records of art catalogs

- ESTC, the *Eighteenth Century Short Title Catalog* containing the bibliographic records of 18th Century titles in Britain

- *RLG Conspectus Online*, a collection development tool for analyzing collections and formulating future collection policies of the member institutions.

Technical services RLIN technical services are only available on the RLIN private network. Input and update of records, set processing, BRCON, copy cataloguing, and other technical services must be done via a dedicated line with a RLIN account, password, and appropriate software supplied by RLIN (Chapman, 1990). Again, this is a matter of policy, rather than a technical decision.

4.1.2 Connectivity projects in Canada

The status of library/research network connectivity in Canada is very dynamic. Many of the mid-level provincial research networks that connect to CAnet, the Canadian TCP/IP backbone network, have come online within the last year with more expected in the near future. While there is little published information available about library activities involving these networks, preliminary data from a recent National Library of Canada survey of CARL libraries (Canadian Association of Research Libraries) and CASUL libraries (Canadian Association of Small University Libraries) indicate that approximately 82% of these libraries have access to CAnet through their respective provincial networks.

Network-accessible information resources Similar to patterns on the *Internet*, online public access catalogues (OPACs) are the most common form of information resource available on CAnet. Fifty five percent of sampled libraries currently offer access to their OPACs, and another 27% have indicated formal plans to do so. Twenty-one percent offer access to citation databases, while another 24% offer access to an on order file. Other information resources listed by respondents include a university information database, an archives database, the *Materials Safety Data Sheet Database*, a government publications database, and a film/video database.

Electronic document delivery An interesting finding of the survey was that 12% of the sample indicated that they are currently using CAnet for transmitting telefacsimile images, while another 21% have stated that they have formal plans for doing so. While no libraries presently use the network for sending scanned images, 18% of sampled libraries have indicated that they have plans to do so.

4.2 Library and research network projects in Europe

4.2.1 The United Kingdom

The United Kingdom's major research network is JANET, a network based on a homogeneous set of protocols. The long-haul network layer is X.25, while higher layers are based on the Coloured Book suite of protocols (Quarterman, 1990). The network connects the local area networks of many academic and research institutions in the U.K. British libraries have been using services across JANET since 1985.

4.2.1.1 Library services on JANET

The integration of library applications into research networks appears to be further advanced in Britain than it is in the United States. For example, more progress has been made on JANET in the supply of technical services over the network; all major bibliographic utilities have linked, or are in the process of linking, their databases to JANET (Stone, 1990b). While most bibliographic utilities are connected, the methods of data transfer tend to be interactive and system-specific, and

do not use standard network facilities. The CURL cataloguing database does, however, offer to deliver records by NIFTP file transfer (between Amdahl and Geac, for example), and BIDS (Bath ISI Database Service) routinely uses electronic mail to return the results of searches. Another advance is that JANET provides a consolidated communications channel for most academic libraries. Hitherto, separate channels would have had to have been maintained for online services (a modem for calls to the public data network), shared cataloguing (a separate leased line to a utility), document supply (Telex), and loans between academic libraries (electronic mail) (Stone, 1991). JANET now offers access to these services either directly through the network, or indirectly through gateways to PDNs. Other efforts to integrate library functions into JANET include providing network access to commercial journal article indexing and abstracting databases (Information World Review, 1991).

Library systems began connecting to JANET before 1985 with over 20 OPACs recorded during that year (Stone, 1991). By 1988, more than 30 libraries had made their catalogues available through the network (Buxton, 1988). As of March, 1991, the number had risen to 60 catalogues (Stone, 1991). This represents the majority of university libraries in the U.K., a number of polytechnic and college libraries, and several Research Council libraries, and the National Library of Wales (MacColl, 1990). Because access to JANET-accessible libraries and other information resources is based solely on remote login, however, there are complications with addressing, login procedures, and command language variations similar to those in the U.S.

An overview of current state of libraries on JANET is provided in the following section. Detailed information regarding access to and content of library catalogues on JANET is available in a document produced by JUGL (JANET User Group for Libraries) (1991).

Access

Directories JANET does not yet have a formal directory structure to provide information and pointers to networked information resources. Discovering what databases are available and where they are available can be difficult. Potential users of JANET resources do have one advantage over *Internet* users, however, in that information about JANET libraries is available in a centralized source with the NRS name of *janet.news* (numeric: 000050005002). This news service is constructed as a typical bulletin board, containing directories of files, accompanied by a viewing facility, from which users can select items to download to their local computers (Buxton, 1988). It is maintained by the Joint Network Team as a guide to the framework of JANET: service names and addresses, official contacts, guides and directories to gateways of other networks, formal reports, announcements, and the minutes of groups (Stone, 1991).

Most germane to library users is the LIBRARY area on the news service maintained by the JANET User Group for Libraries (JUGL). It contains a file, JANET-OPACS, which contains a list of JANET-accessible catalogues and their addresses, in addition to brief catalogue content descriptions and login procedures. Other files available include a bibliography on JANET use, and a directory of library and staff email addresses. Three mail distribution services support the U.K. library community: LIS-Enquiries for general correspondence, LIS-Info distributing JANET-OPACS and other directories, and LIS-IT for information technology related issues.

Currently, online directory services for the JANET network are under development. The National Information on Software and Services (NISS) provides another major bulletin board, largely contributed by individual computing centres, the Combined Higher Education Software Team (CHEST), and many subject-related projects (Stone, 1991). All U.K. universities are currently installing X.500 directory services. Extending them as guides to resources beyond the primary function of supporting electronic mail has also been discussed (Stone, 1990c).

Connection Interactive connection to hosts on JANET uses the Triple-X (specified in the Green Book), and is made through a software or hardware PAD (Packet Assembler-Disassembler, a device which allows terminals not suitably equipped to connect to a packet-switched network) then making a CALL to a JANET address. Although the underlying service address is based on a 12 or 14 digit number, most PADs can store some or all alphabetic mnemonics provided by the Name Registration Scheme (NRS) to represent the number. The Name Lookup Protocol (NLP), introduced in 1990, allows PADs to update themselves by interrogating automatically a campus or national Name Server, should the user request an unstored mnemonic name (Stone, 1991).

Login procedures Generally, passwords and user IDs are not required to log onto library catalogues on JANET. In a few cases, other information may be required, for example, typing LIBRARY to a SERVICE? prompt, or selecting terminal types. Passwords, user IDs, and other login information is published in the *library.catalogs* file. Most systems will support dumb, scroll-mode, terminals. Many also support VT100 (Stone, 1991).

Search systems and interfaces The range of library systems connected to JANET are those produced by familiar vendors. A sample includes Dynix, Libertas, Geac, Microdata, CATS, Dobis/Libis, and BLCMP. Common access points in the systems are author, title, ISBN, location and classmark. Others systems support these access points as well as publisher, date, and language. Most systems do not allow keyword searching.

The wide range of systems linked to JANET result in the same problem of diversity of interfaces that is found on the *Internet*. While help screens can sometimes mitigate this problem, dealing with new interface every time a new catalogue is searched can be frustrating. Stone (1990c) has noted that JANET use is currently not very high and cites non-standard interfaces and search syntax variations as part of the problem.

The functions available on JANET-accessible catalogues are also limited. Searchers using the online catalogues can only look up material in the catalogues, see whether the book is in library, on loan, or contained in a nearby collection. As yet, the user cannot often reserve a book, or send a request for it online; it must be requested via normal ILL channels (Stone, 1990a).

Information resources

Online catalogues The main information resources available on JANET are the records of the collections of universities, polytechnics, and colleges. The scope of the catalogues vary

according to the research interests of the local institution as well as the local policies concerned with what is to be included in an online catalogue. For example, some systems do not include serials at all, while others post them as titles or include individual volumes as if they were monographs. Access to serials in a few other systems, such as Edinburgh's EULOLIS, is provided through an online union list of serials. The items that are represented in the catalogues are generally those found in OPACs, such as monographs, dissertations, and maps (Buxton, 1988).

Catalogue coverage of materials held in collections is as varied as it is in the U.S. While many are almost complete, some can be as low as 10 percent (e.g. University of Cambridge). While this may be indicative of the inherent difficulties of retrospective conversion, the fact remains that records that are in non-electronic form are not available to users doing their searches over research networks.

Miscellaneous information items also available in the networked catalogues include recent accessions, reader's comments, user loans and reserves, information about computing services, campus telephone directories, bus and train timetables, job boards, and entertainment information (Buxton, 1988). Such information is also widely found in the campus bulletin boards and information systems provided by many Computing Centres (Stone, 1991).

Non-catalogue information resources A number of bibliographic and other databases are available on JANET in addition to the catalogues. Examples include:

- BIBLIO, the Birkbeck Computer Literature Database
- Southampton Library's collection of the papers of the first Duke of Wellington. Some are available in full text
- Numeric databases including portions of the *Cambridge Structural Databank*, *Inorganic Crystal Structure Datafile*, and the *Crystal Data Identification File*
- *Fine Chemicals Directory*, *Thermodynamic Data*, *Interactive Molecular Graphics Package*, and *Crystal Structures of Metals and Alloys*
- *Time Series Databank*, a macroeconomics numeric database
- CRONOS ICG *Databank*, contains data for the Statistical Office of European Communities
- NOMIS, the National Online Manpower Information System containing employment, unemployment, vacancy and population data in Britain
- AGDEX, the Edinburgh School of Agricultural Database of grey literature on applied agriculture.

The above databases are provided free of charge by various institutions with the exception of NOMIS and AGDEX, which require payment for connect time and, in some cases, royalties for items downloaded (Buxton, 1988). The number of such nationally available services is likely to increase following a report of a sub-committee, the National Datasets Task Group, to the Computer Board, the main funding agency for academic computing in the U.K. (Stone, 1991).

Another major bibliographic resource available through JANET is Blaise-Line operated by the British Library. Blaise-Line provides access to a number of databases for both online searching and downloading of MARC records. These databases include:

- BNB MARC, containing MARC records of the *British National Bibliography*
- LC MARC, containing MARC records from the Library of Congress
- *Whitaker File*, an electronic version of *British Books in Print*

While there are obvious advantages to connecting these technical services to JANET, full utility of the research network is limited by the fact that neither the bibliographic suppliers' systems, nor those of the client libraries, are based on Coloured Book or OSI protocols (Buxton, 1988). Document requests and bibliographic records, therefore, cannot be sent directly to a client library's database through file transfer. Full integration of these systems into the research network would require the adoption of common standards.

Licensed databases Beginning January, 1991, users at all subscribing U.K. higher educational institutions connected to JANET nationwide will be able to search *ad lib* a variety of databases from the Institute for Scientific Information (ISI). The Combined Higher Education Software Team (CHEST) has leased four ISI files for an annual flat rate fee to allow end-users unlimited searching. The database will reside on a host at Bath University Computing Services and will be made available to the institutions involved via JANET (Information World Review, 1991). Results are delivered by electronic mail to the searcher's own mailbox within minutes (Stone, 1991).

The ISI databases include:

- *Science Citation Index*
- *Social Science Citation Index*
- *Arts and Humanities Citations Index*
- *Index to Scientific and Technical Proceedings.*

Gateways to online services In addition to the above information sources available directly over JANET, the network also provides access to commercial database vendors. This is accomplished in a number of ways:

- ***directly through JANET***. At present, the only suppliers of commercial database services directly connected to JANET are the British Library and OCLC, but with the relaxation of the criteria for connection and the growth of end-user searching, this is expected to increase.

- ***through gateways to other research networks***. Links to other networks is accomplished through the IXI backbone, which now links all European networks. In this way, STN (a Frankfurt-based host for *Chemical Abstracts*, INSPEC, and many other services) connected to the German network DFN, is now accessible at no communications cost to many users in the U.K. and elsewhere.

- ***through a gateway to the British Telecom PSS service (the "Global Network Service")***. The gateway between JANET and PSS represents a considerable cost savings and a reduction of information channels; no telephone lines with their concomitant noise problems or modems are required, the data exchange rates are higher, and libraries can avoid individual subscription charges to the PDNs. Based on standard charges, overall cost savings on telecommunications through JANET gateways can be as high as 54% compared with dial-up access (Buxton, 1988). Access to PDNs is restricted to registered users with a charge incurred for usage (though several hosts (DIALOG and ESA, for example) permit reverse-charging where the users' PSS account is not debited directly) (Stone, 1991). The JANET Network Executive has set up accounts and passwords for individuals for use of the JANET-PDN gateways. Once a given allocation is exhausted, the system refuses access to that user ID.

Availability There are no restrictions for use of the university catalogues available on JANET. Users may need to be registered (at no charge) to use some of the more specialized services. The Bath ISI Database Service (BIDS) is available only to staff and students of subscribing institutions within the U.K (Stone, 1991).

Interlibrary loan and document delivery

A large majority of ILL requests in Britain are handled centrally. Fully 90 per cent of all requests are sent through the British Library's Document Supply Centre (BLDSC) (Stone, 1990a). The facility which handles the requests is ARTTel, an online service for requesting loans or photocopies from the BLDSC near Boston Spa, West Yorkshire (Buxton, 1988). ARTTel, previously available through dial-up and PDNs, linked to JANET in 1987. Documents are requested by creating a text file of the requests and uploading them to the ARTTel system. Use of ARTTel through JANET, however, remains low. Of 40 percent of requests that BLDSC receives from academic libraries, only eight percent route through ARTTel (Stone, 1990c). Reasons cited for its low use include the clumsy way in which ARTTel files are created, edited, and sent to the system, and a lack of awareness that the system is available through JANET.

The fact that most university catalogues are accessible through JANET, and that all libraries now have email addresses, may change the way in which ILL in done in the U.K. More than 60 libraries have created separate email addresses for ILL requests and a interlibrary loan mail

distribution list has been created. Many university libraries in Holland, France, and the U.K. will contribute to an initiative called Project ION for the development of interlending using OSI protocols over X.400. The U.K. participation in this project is organized by LASER (Stone, 1991).

Technical services

As in the U.S., libraries in the U.K. have been organized into library networks for the purpose of sharing resources and bibliographic records. In contrast to American bibliographic utilities, however, U.K. utilities offer their technical services over JANET. The utilities had been based previously on central databases accessible through private leased lines or PDNs. As early as 1987, however, they began offering their services on JANET (Buxton, 1988). Currently, all major utilities in the U.K. have linked or are in the process of linking their databases to JANET. Basic subscription and display charges still apply, however, when connecting to some bibliographic suppliers such as the British Library (Stone, 1990b).

A major bibliographic resource sharing scheme on JANET has been organized by CURL, the Consortium of Research Libraries. This project was originally funded by the University Grants Committee (UGC) designed to enhance library automation through JANET by making available, free of charge, a database of the bibliographic records of CURL libraries (Heaney, 1989). Members of CURL are the Universities of Cambridge, Edinburgh, Glasgow, Leeds, London, Manchester, and Oxford. This scheme is similar to those in North America in which members deposit their own records in standard MARC format making a pool available for shared cataloguing. Functionality is limited, however, because the systems involved lack sophisticated communications software that support direct file transfer of records. Records are instead copied by simple screen capture or by making batch requests for tape transfer (Stone, 1990c). The services, however, represent a major step forward for libraries using the research network (Stone, 1990a).

Professional activity

The development of network services by librarians in the U.K. has been encouraged by the active support of the Joint Network Team, responsible for the management of the national network. Their initiative in 1986 led to the creation of the JANET User Group for Libraries (JUGL). All libraries of institutions eligible to connect are automatically members who elect a committee to promote library interests in networking. JUGL has worked to stimulate a wider awareness of the potential of the library contributions to academic networking, as well as a readiness to develop and adapt professional services.

Among a variety of JUGL activities to date worth noting are:

- the JUPITER Project: a two-year programme in networking for university libraries, funded by the UGC
- the U.K. office for Library Networking, funded by the British Library

- the JUGL Newsletter, produced by Brunel University
- JANET-OPACs and the *Directory of email addresses for libraries, services and staff*, both produced by the University of Sussex
- annual conferences and regular workshops

Changes expected in 1991

In 1991, a new central Government-funded agency, the Information Services Committee (ISC) of the Universities Funding Council (UFC) will take over the oversight of all aspects of computing in U.K. universities, to include for the first time the needs and services of libraries and university administrations. This will extend the long-standing coordination of research and academic computing (and of network development and standards) provided by the Computer Board for Universities and the Research Councils. Simultaneously, the management of JANET will be restructured within the Networking Association. These major changes offer considerable opportunities for the development of library networking in the U.K. (Stone, 1991).

4.2.2 Continental Europe

The major library and research network projects in the remainder of Europe have taken a more integrated approach to the use of research networks. Many libraries in continental Europe have taken steps to fully integrate their systems into the infrastructures provided by their respective national networks. In Finland, for example, where libraries are just beginning to automate, a project is underway to integrate, from the outset, a homogeneous set of library systems into the Finnish research network (Soini, 1990). In the Netherlands, an existing bibliographic network will soon be integrated into SURFnet, the national Dutch research network, at which time connectivity for all library work will be supported by the research network (Costers, 1990). A similar pattern emerges in Sweden where all university libraries are to be connected to SUNET, along with their library network, LIBRIS (Olausson, 1991). In these countries, the pattern is for the most part not *ad hoc* connections of library systems, but a more full integration of systems into existing research network infrastructures.

4.2.2.1 Sweden

There is currently a national project underway in Sweden to connect library systems to the Swedish national research network, SUNET. SUNET connects universities in Sweden through the connection of local Ethernets to national Ethernets via 64Kbps lines. TCP/IP, DECNET, and NJE EARN logical networks are all carried over the SUNET backbone. It is intended that both academic libraries and *LIBRIS*, Sweden's national bibliographic database, will use the network for their communication needs (Olausson, 1990).

Access

An X.500 directory pilot project is currently in place to provide pointers to information resources and electronic mail addresses.

Access to catalogues and other information resources is via remote login using IP or domain name addresses, and are available to all SUNET users with usernames and passwords (Olausson, 1991).

Information resources

There are currently over five university library catalogues connected to SUNET, although all systems located at Swedish academic institutions are to be connected eventually. The range of library systems connected include Dobis/Libis, Libertas, and VTLS, as well as planned connections for Geac and Tinlib systems. Non-OPAC information resources currently include:

- *Medlars*, the medical database covering physiotherapy, occupational therapy, rehabilitations, terminal care and alternative medicine. Use of this database incurs connect-time charges (MIC News, 1990)

- *BASUN*, a database covering the services available on SUNET.

Other resources planned are:

- a Swedish legal database
- a journal article database.

A standardized interface for these resources is also planned (Olausson, 1991).

Technical services

Shared cataloguing is currently supported by LIBRIS, the Swedish national bibliographic database. Availability of this database is anticipated very soon (Olausson, 1991).

4.2.2.2 Norway

Libraries in Norway have also begun to utilize research networks to support library activities. Unlike the separate proprietary library systems used in the U.S. and the U.K., library automation in Norway is based on a centralized cooperative integrated library system called BIBSYS. Beginning in 1972 as a project shared by four Norwegian institutions (The Computing Centre at the University of Trondheim, the National Council for Academic and Special Libraries, the Technical University Library of Norway, and the University Library of Trondheim), BIBSYS currently serves the Norwegian community of universities and academic colleges. The system consists of a centralized database that is used by participating libraries to support subsystems for acquisitions, cataloguing, bibliographic information retrieval, and circulation control. The system is currently based on ADABAS software and runs on an IBM 4381 computer located in Trondheim (BIBSYS, 1990).

BIBSYS began using the national research network in Norway, UNINETT (University Network), in June of 1989 (Gunderson, 1991). UNINETT is designed to connect all higher educational

institutions in Norway and is based on inhouse UNINETT protocols, as well as DECNET and Coloured Book protocols. Standard network services provided are remote login, file transfer, conferencing, mail.

Libraries currently participating in BIBSYS include:

- Norwegian College of Veterinary Medicine, The Library
- The National Library, division Rana
- The Norwegian School of Economics and Business Administration
- The Technical University Library of Norway (NTUB)
- The University Library of Bergen
- The University Library of Oslo
- The University Library of Tromse
- The University Library of Trondheim (UBT)
- The University of Trondheim, the Medical Library
- Trondheim School of Economics and Business Administration.

Access

Directories Guides to information resources available on the network can be found online at the email fileserver *info@uninett.no*, or through anonymous FTP to *anm.uninett.no* (129.241.1.99). BIBSYS and the University of Oslo have printed guides.

Connection Connection to catalogues and databases is accomplished through remote login. Usernames and passwords are required.

Interfaces Interfaces to BIBSYS include PUBSOK, a user-friendly menu-driven program with a limited number of searchable fields, and GENSOK, a command-driven interface based on the Common Command Language (ISO 8777) standard. BIBSYS is also participating in the development of IANI (Intelligent Access to Nordic Information Systems), an effort by the Nordic Council for Scientific Information and Research Libraries to provide a common format for login and access to multiple databases and hosts. This interface is also based on ISO 8777 (BIBSYS, 1990).

Information resources

Information resources available on UNINETT include the BIBSYS database, comprising over one million database records, and several other catalogues and databases managed by the University of Oslo Library (Gunderson, 1991).

The University of Oslo Library's online catalogue contains three databases (University of Oslo Library, 1991):

- **BOK**, a database of Norwegian books containing over 122,000 records

- **NOTA**, a database of selected articles from 400 Norwegian periodicals containing over 88,000 records

- **SAMBOK**, a Union Catalogue of foreign an Norwegian books in 350 Norwegian libraries containing over 640,000 records

The University of Oslo Library also provides access to several databases in TRIP, a Swedish database system produced by PARALOG AB in Stockholm. The TRIP databases include (University of Oslo Library, 1991):

- **BAUER**, a discography of old 78-records from 1900-1950.

- **BIBLIOTEK**, containing addresses of libraries participating in the Nordic Union Catalogue of Periodicals and the Norwegian Union Catalogue

- **EMNE**, Dewey Decimal System index of subject headings to the catalogue at the University of Oslo's (UiO) Faculty of Social Sciences

- **IRIB**, the library catalogue at the Norwegian Research Center for Computers and Law, UiO

- **LC86,-87,-88,-89**, Library of Congress (U.S.) MARC records

- **LOVTITLER**, an index of titles of Norwegian laws

- **LYDOPPTAK**, a discography Norwegian gramophone records, CDs, and cassettes

- **MUSBIB**, an index to articles about music in Norwegian periodicals

- **MUSIKKTRYKK**, containing Norwegian musical prints at the National Collection of Music at The University Library Of Oslo (UBO)

- **NOSP**, Nordic Union Catalogue of periodicals in Danish, Finnish, Icelandic and Swedish libraries

- **PALSAM**, a database of paleotypes in the The University Library Of Oslo (UBO) collection (under development)

- **RARLIB**, a database of rare books at The University Library Of Oslo (UBO) (under development)

- **SAMPER**, Norwegian Union Catalogue of Periodicals

- **SANGINDEX**, index of titles and first lines of Norwegian songs from the National Collection of Music at The University Library Of Oslo (UBO)

- **TEATERHIST**, Norwegian theater history (under development)

- **TX**, Norwegian periodicals at the Norwegian Department of THE UNIVERSITY LIBRARY OF OSLO (UBO)

- **UTMUSIKK**, foreign musical prints from the Catalogue of the National Collection of Music at THE UNIVERSITY LIBRARY OF OSLO (UBO).

Technical services

The technical service functions supported by BIBSYS are currently available over UNINETT, or are under development. Specifically, shared cataloguing is available on the research network via the centralized BIBSYS database. Exchange of acquisitions information is planned for BIBSYS and will be tested this year. UNINETT is also presently testing a FAX transmission service over the network to be used for electronic document delivery (Gunderson, 1991).

4.2.2.3 Finland

In Finland, a project has been underway since 1988 to connect university library systems to the nation research network, FUNET. FUNET is a TCP/IP-based network connecting over 36 universities and research institutions. As in the Netherlands, the plan is to use the infrastructure of the national research network to connect libraries for bibliographic information provision, document supply, and technical services. This logical library network is called LINNEA.

LINNEA LINNEA (Library Information Network for Academic Libraries) is a proposed logical information network that is the result of a coordinated plan by the Finnish Ministry of Education. Finland has maintained a central union catalogue since 1978. A movement toward creating local library systems in Finnish libraries in the early 80's, and the perceived threat to library cooperation that this entailed, inspired Finland's Ministry of Education to develop a plan to coordinate the automation of Finnish libraries. The plan involves first establishing library systems at universities based on the same software, then providing connections both among local systems and between those systems and the central database (Soini, 1989).

The first stage in the plan has been largely completed. VTLS (Virginia Tech Library System) software was chosen in 1988 for installation at the university libraries. The first system became operational at Lapland University in 1989; the remaining systems were to be installed by the end of that same year.

LINNEA and FUNET Once the systems are installed, the second stage of the plan is their interconnection. Connectivity options provided by the VTLS software are either machine-to-machine connection using dial-up, leased lines, or network-to-network links. Because there is already a well-developed backbone network connecting university LANs, namely FUNET, the

latter network-to-network option was chosen. In addition, the use of FUNET will provide the libraries with the enhanced network services supported by the network. Once LINNEA is fully implemented at the end of 1993, it will provide users of FUNET access to both the central bibliographic database and other library services throughout Finland (Soini, 1989).

4.2.2.4 The Netherlands

There is an ambitious project underway in the Netherlands to connect all Dutch libraries to the Dutch national research network in a full OSI implemention. The major research network in the Netherlands is SURFnet, a value-added network connecting Dutch universities, and governmental and private research centres. The network has recently switched to X.25 protocols and has implemented OSI standards. Pica, a central bibliographic database, and other local library systems, will be integrated into SURFnet in the near future (Buitelaar, 1990).

The Pica database Pica, or Project for Integrated Catalogue Automation, is the Dutch equivalent of a centralized bibliographic resource such as SWALCAP in the U.K. It began in the early 1970's when major research libraries in the Netherlands participated in the creation of a shared catalogue. At the present time, it contains over six million titles.

The Pica network Early in its development, Pica established a private bibliographic network to link local library systems at universities with the central database to support shared cataloguing and interlibrary loan. In 1985, the "Picalink" project enhanced these centre-to-local connections with an X.25 network based on the OSI model. Acquisitions support was added at this stage (Dekkers, 1990). It is also experimenting with Virtual Terminal Services to provide local-to-local access to applications running on other library systems, in addition to access to the central database.

Pica and SURFnet SURFnet and Pica recently signed an agreement to make Pica's library network part of the SURFnet infrastructure. Pica's existing network, based on leased lines, will be integrated into SURFnet's X.25 network. Dutch universities will be the first to be connected, followed by institutions engaged in research, and public libraries. Libraries in the Netherlands will receive expanded services as a result of this integration; SURFnet services, based on OSI protocols, will be remote login (VTP), file transfer (FTAM), email (X.400), among others. These services will be available to Pica libraries in addition to those previously provided by Pica (Buitelaar, 1990).

SURFnet users will also benefit from the incorporation of the Pica network. All SURFnet users will be able to access the central bibliographic database as well as those situated at the member institutes. Because SURFnet itself is connected to other European and international networks, access to Pica information resources will eventually be available outside the Netherlands. These services will also be based on OSI protocols, namely, the ILL protocol, and Search Retrieve (SR).

4.3 Library and research network projects in Australia

Australia has recently inaugurated a new research network, called AARNet (Australian Academic and Research Network), which is at the focus of library and research network activities in this country. Commissioned in May, 1990 and based on TCP/IP and DECNET protocols, this network replaces SPEARNET and ACSnet, two early Australian research networks. At present, there are 33 academic and research institutions connected to AARNet located all across the Australian continent.

4.3.1 Library systems on AARNet

Although AARNet has only been up and running since the spring of 1990, over half of the institutions connected to AARNet have made their library systems available. By October, 1990, 17 of these 33 institutions had connected their libraries. Despite this fast-paced development, the full potential of the research network, as with many of the networks discussed in this chapter, remains to be developed.

Access

The only directory of AARNet library resources is currently a document entitled, *AARNet Resource Guide* (AARNet, 1990). It contains a section providing the addresses, login procedures, and information resources available on the AARNet-accessible library systems. Access to the library systems is accomplished through TELNET, with domain name addresses similar to those used on the *Internet*. For example, the Baillieu Library at the University of Melbourne is available by connecting to *library.unimelb.edu.au*. These connections remain somewhat clumsy, however, similar to other library connections based on remote login. The terminal emulation needed for connection is most commonly VT100.

Information resources

The primary form of information resource available on AARNet, as on the *Internet* and JANET, is the library catalogue. The catalogues currently connected represent the collections of 17 universities, colleges, institutes of technology, as well as defence academy. Notable collections available include those in military history, Australiana, Canadian studies, health sciences, Latin American studies, nursing, and romance literature. Media include books, periodicals, maps, films, videos, audio-cassettes and music scores.

Other information resources include:

- the ***Mihilist Database***, covering Australian military history
- the ***Austlit Database***, covering Australian literature.

The catalogues are available to all registered AARNet and *Internet* users. The specialized databases are currently available free of charge (AARNet, 1990).

Technical services

It is unclear at this time whether technical services are, or are planned to be, supported over AARNet along with end-user information services, however, the Network Technical Manager of AARNet, Geoff Huston, recognizes the network's potential in this area. He states that "the requirements of libraries in terms of data communications services do not stop at remote interactive access", but also include "the support of interlibrary loan applications, database-level catalogue access, distributed catalogue access, transmission of digitally encoded material... ...and the support of electronically published material" (Huston, 1990a:11). He also states that the realization of these applications will be largely the responsibility of the library sector.

4.4 Library and research network projects in Japan

In Japan, two major academic and research networks connect researchers and scholars; only one, however, is involved in library linking projects. This network, commonly called NACSIS, has evolved from an academic network providing network communications and supercomputer access into an integrated computing and information resource. Japanese libraries have been integrated into NACSIS since 1987 and in this way follow more closely the pattern emerging in Europe. As early as 1986, plans were in place to expand Japan's existing research network to include connections to academic libraries. This network has a high level of integration, supporting both access to information and technical services. The other major research network in Japan is JUNET, a network that functions as a testbed for network experimentation and, consequently, has no connected libraries.

4.4.1 NACSIS and the Science Information Network

NACSIS, or the National Centre for Scientific Information System, is primarily an independent inter-university research institute and bibliographic utility (Naito, 1989). The NACSIS organization was created in 1986 out of the Centre for Bibliographic Information based at the University of Tokyo, an older body that had previously developed an online shared cataloguing system. In 1986, NACSIS joined with Japan's first academic network, N1, to create the *Scientific Information Network*. Its stated function was not only to connect academic computing centres in Japan, but to connect academic libraries as well. The acronym, NACSIS, is commonly used to denote the network (Quarterman, 1990).

N1, Japan's first inter-university network, was established in the early 1980's and was joined in 1986 by NACSIS. A year later, NACSIS began to expand N1 using high-speed digital lines rented from Nippon Telegraph and Telephone, Japan's PDN (Ishida, 1988). In an effort to create the *Scientific Information Network*, it offered packet-switching services based on these lines to university libraries, computing centres, and research institutes without charge. Services provided

by the network to these institutions, based on N1 protocols, are remote time sharing, remote job entry, file transfer, and electronic mail.

Access

Access to the library resources is through a remote connection.

Information resources

The union catalogue contains bibliographic records and holdings files of monographs (two million records) and serials (1.9 million records). This database is available for searching through the NACSIS network.

NACSIS also provides access to a number of commercial databases. No details, however, are available regarding licensing agreements or charging structures for use of the databases. Example databases include:

- CAS, a chemical abstracts database
- MOL, a molecular data file
- XDC, a database of crystal data
- INSPEC, the business database
- JMARC, a database of Japanese books
- METADEX, a database of metals
- PROTEIN, a protein databank
- GEODAS, a database geology data
- COMPENDEX
- SCI
- EMBASE
- *Harvard Business Review* in full text.

NACSIS also produces inhouse databases, including abstracts of Ministry of Education funding research, and an index of doctoral dissertations granted by Japanese universities (Naito, 1989).

Technical services

The national union catalogue maintain by NACSIS is available to cataloguers over the NACSIS network. Cataloguers can access holdings, as well as authority records, and do copy cataloguing from files such as US MARC, UK MARC, Japan MARC, and Book Traders MARC. The cataloguing system itself is capable of processing Kanji characters. Interface software for the system is located in the local computer providing full screen interaction (Naito, 1989).

It is apparent from this overview of library/research network connectivity projects that the utilization of research networks for library activities around the world is at a very early stage. While many information resources are beginning to become available, they remain rather difficult to use. It is difficult to find what information resources are available, to search them with consistent interfaces, and to directly download search results. Other advanced services, such as electronic document delivery, are as yet virtually non-existent. Further, in areas where extensive

bibliographic and library networks have developed to support technical services, such as in the U.S., these services have not been integrated into the infrastructures of research networks. While there are intentions among British, European, and Japanese libraries to use research networks for technical services, the network services to support these activities have yet to be fully developed.

5. ISSUES IN LIBRARY USE OF RESEARCH NETWORKS

Although the end-user applications discussed in the previous chapter have the potential to form, in part, the basis of the rich information environment that will be available on research networks, there are still numerous issues to be resolved before these applications can be fully implemented. They include both technical issues involved in the development of the actual applications as well as policy issues that will arise as research networks change the established patterns, processes, and organizations associated with information production, access, and distribution. This chapter will outline and discuss some of the major technical and policy issues that will have to be overcome in the process of creating information networks.

5.1 Technical issues

5.1.1 Library system/research network connectivity

With the present generation of library systems, the difficulty of simply connecting to the research networks presents the greatest immediate impediment to the more widespread use of the facilities offered by such networks. Giving a user access to the resources available via a research network is not as simple as establishing an account with OCLC. The first step is to have access to a computer host that is connected to a research network. Unfortunately, the computer which is central to library operations, the integrated library system, often cannot connect to anything. This isolation is due to that fact that most of the current generation of library systems are closed systems; they use master-slave architecture (a central CPU supporting a number of terminals) and proprietary communications protocols. They do not, in effect, support the peer-to-peer communication that is characteristic of research networks.

In this closed environment, the library wishing to make the facilities of a research network available to its patrons can generally only do so by acquiring a separate "network workstation" that connects to the network. This would require a significant investment. Even if a workstation or personal computer is already available, there will be the need to acquire additional hardware, such as communications co-processors, as well as software to implement the network protocols. Given the start-up costs, many libraries may come to the conclusion that the benefits available from the networks at the present time do not justify the expense. This situation is of course a classic chicken-and-egg dilemma: librarians are unwilling to go to the expense of connecting because the benefits are marginal, while developers are unwilling to go to the expense of improving services over the network because the user base is too small. (Of course, if the library

is already part of the "wired campus", the problems are different; the question changes from one of providing simple access to one of integrating remote services with local services.)

The problem would not arise if connecting to the network involved little or no additional expense. One means of reducing additional costs would be to ensure that connectivity came as part of equipment the librarian is buying anyway, such as a new library system. This is the approach adopted by the US Defense Advanced Research Projects Agency in funding the integration of TCP/IP into a version of the UNIX operating system; all users who acquired that or subsequent versions of UNIX automatically gained the protocol suite needed to connect to the *Internet*. Because approximately 90% of American university computer science departments had one or more UNIX systems, there was enough of a "critical mass" of users that implementations of TCP/IP became widespread (Comer, 1988). In the library world, however, such a move would be more difficult to achieve because of the wide range of systems that are currently in use in libraries. Integration of the protocols into the most common library system, for example, would achieve penetration of at best 29% of large Canadian libraries (Merilees 1989).

Neither is the problem of system connectivity as "simple" as integrating the protocols into one or more operating systems. Although there are a number of library systems which use the UNIX operating system (e.g. Inlex, Minisis, multiLIS, VTLS), very few of these offer the connectivity inherent in the operating system. This stems from the fact that, while the operating system supports network connection, the library applications do not. The applications, like the vast majority of the current generation of library systems, are "closed" rather than open.

The problem of closed systems will not be solved either quickly or easily. System vendors will not undertake the expensive modifications necessary to achieve conformance with networking protocols without some confidence that their investment will be rewarded. At the same time library administrators will be reluctant to embark on major system changes without clear evidence that the changes will save money or are essential to meet the information requirements of their patrons.

The problem of connectivity also needs to be solved at several levels in order to make meaningful interoperation of disparate systems possible. Attention to connectivity must take place at the level of:

- the computer/operating system supplier
- the application software supplier
- the data used within applications.

Each of these aspects will be discussed further in the following sections.

5.1.1.1 Computer/Operating System Connectivity

The first problem at this level is to simply enable the computer to connect to the network. This is often surprisingly difficult to do with the current generation of library systems. Even medium-

sized and large computer systems often cannot readily support host-host communications (for example IBM CICS and the Pick operating system).

It should be noted that this is a difficulty most library system vendors cannot overcome on their own. Most vendors chose the hardware/software platforms on which their systems have been built for reasons other than connectivity and are now committed to those platforms. They are dependent, therefore, on the computer manufacturers to add appropriate connectivity to existing platforms.

Fortunately, almost all of the major computer manufacturers, under pressure from government purchasing programmes, have implementations of TCP/IP and have further made a commitment to implement the OSI protocol suite. Many of the most widely implemented library systems, however, will continue for some time to use operating software which is inherently not amenable to network connection.

5.1.1.2 Application Connectivity

The largest connectivity issue to be overcome arises at the level of the library application software, such as an online system. Even when the computer system can connect to the network, the user frequently cannot connect to a remote library application because the application has not been designed to be accessed by this method. Even less likely is that the user will be able to connect to a remote system from within the local library application, such as a cataloguing system, public catalogue, or any other library function.

The most rudimentary kind of connection to a library application is through use of remote login such as TELNET on TCP/IP networks, and Virtual Terminal within OSI. Supporting remote login to a catalogue raises numerous questions of application and program design, such as supporting generic rather than specific terminals, permitting explicit logging on and off, designing screens and help messages which do not rely on the capabilities of specific hardware (e.g. colour, specific function keys) (Lynch, 1989b). Supporting more sophisticated kinds of interconnection from within applications, such as automatic redirection of failed searches from the local catalogue to a remote database, is much more difficult. This may require the complete redesign of library applications to support the kind of distributed systems which are implicit in protocols such as file transfer, remote database access, Z39.50.

Modifying applications to support connectivity may be a less serious problem for the large academic libraries, which generally have sizable mainframe systems, than for smaller libraries. While large mainframe systems may also have difficulty supporting peer-to-peer communications, they have the advantage of being supported by specialist systems staff who are often capable of devising and implementing appropriate front ends and gateways. This is not generally true of medium and small library systems. Usually, there are no local staff with the background to undertake such changes. It may be up to system vendors to change these smaller systems. It is significant, however, that vendors of small to medium sized library systems have demonstrated very little interest in open systems or application connectivity.

Librarians will be unable to make efficient use of the resources available on the network until they can connect from within their application or until their application can be accessed from the network.

5.1.1.3 Data structures

The third level at which the problem of connectivity must be addressed is that of data structures. Librarians are well accustomed to using standardized data structures. The MARC record for monographs, for example, has been standardized for over 20 years now, and this was itself based upon the structure of the Library of Congress catalogue card distributed in a standard format since 1907. Outside the world of library operations and systems, standardization of data structures is much less advanced. There are, for instance, numerous different data structures used by abstracting agencies or by indexing agencies; providing a library user with integrated or consistent access to these will form a significant challenge.

The problem of lack of standardization of data structures becomes more severe when providing electronic access to documents rather than information about documents. There are a large variety of proprietary word processing formats for textual material as well as graphic formats for image data. Furthermore, a single document may contain a mixture of text and images. While many of these formats can be converted from one to another, a loss of information is often a consequence.

There is a considerable amount of work taking place in the standards community to develop standards to allow the exchange of both textual and graphic structured information. Among the standards which the librarian should be aware of are *Standard Generalized Markup Language* (SGML), *Office Document Architecture* (ODA), and *Page Description Languages* (PDLs).

5.1.5 Transition to OSI

This section deals almost exclusively with the transition from TCP/IP to OSI. This is because, of the three most widely-used protocol suites (SNA, DECNET and TCP/IP), TCP/IP alone is non-proprietary and therefore issues surrounding it are more readily debated in public forums. Nevertheless, the issues of co-existence with and transition to OSI are similar for all the major protocol suites, although the implementation details will obviously vary.

Note that co-existence is not the same as transition: *transition* implies a community consisting of networks that currently contain only one protocol suite (i.e. TCP/IP) that wishes ultimately to contain only another protocol suite (i.e. OSI). *Coexistence*, on the other hand, implies two communities, each containing a different protocol suite, which wish to interoperate with each other. The approach taken here is that a strategy which allows co-existence will eventually lead to transition as one of the protocol suites "withers away" for lack of support, or as the product of the union between OSI and TCP/IP reaches maturity.

Interoperation can be accomplished in several ways: the dual stack approach, the protocol gateway approach, and the service bridge approach.

- ***Dual stack approach*** In the dual stack approach, some, or all, computer systems in a network implement both protocol suites, at least as far as the data link level where the protocols become identical. The most effective approach requires both suites to be present in all nodes of the network (*dual-stack* nodes). In practice, however, many nodes of the network will contain only one protocol suite (*uni-stack* nodes). This heterogeneity can lead to difficulties when different stack types attempt to communicate. For example, communication between two dual-stack nodes or between dual-stack/uni-stack pairs can take place because the presence of a protocol suite common to them both is guaranteed. Difficulties can arise, however, when the user on a dual-stack system does not know which protocol suite to use when communicating with any particular uni-stack system ("Do I use FTAM or FTP?"). Further, the user may not know in the first instance that the system with which he/she is communicating only supports one stack. A more serious problem results when two uni-stack nodes try to communicate; one uni-stack node will be able to communicate with another uni-stack node only if the protocol suite they support is the same.

- ***Protocol gateway approach*** In this approach, a number of dual-stack systems are present in the network that can translate between one protocol suite and the other; the dual-stack node acts as a *gateway* between the two protocol suites. This approach is at present widely implemented for electronic mail applications. It could also be implemented for other network applications which are functionally similar in the two protocol suites.

  There are, however, a number of problems involved in using gateways to interconnect incompatible networks. The processing required at the gateway can affect performance severely. In addition, if there are only a few gateways available, there is a substantial risk of bottlenecks developing. The services offered by the two protocols will inevitably be different, and a mapping between the two will necessarily result in a loss of information. This loss, at its least, can result in the annoyance of a user and, at its worst, to catastrophic situations, such as a routing loop. This situation is worsened when two or more gateways need to be traversed. Gateways should therefore be seen primarily as short-term solutions to achieve connectivity, rather than long-term means of achieving co-existence.

- ***Service bridge approach*** The service bridge approach involves devising mechanisms to allow the services of one protocol suite to be available via the other. In this case, to allow TCP/IP networks to support OSI upper layer protocols, and thus OSI applications.

TCP/IP is by far the most commonly used protocol suite for interconnecting peer-to-peer computer networks in the academic and research community. It is now so well accepted as a proven, stable, and mature protocol suite that it will undoubtedly remain viable for many years to come. Sales of technology based on the TCP/IP protocol suite continue to grow, despite the public commitments of national governments and major vendors to OSI. This two path approach

is particularly true of Europe, even though Europe has been a strong supporter of OSI in the international community. Nevertheless, most experts in the area agree that ultimately networking based on OSI application standards will supersede TCP/IP (DiDio 1989; Fisher 1989; Rose 1990). This will take place because of government regulation, through GOSIP (Government Open Systems Interconnection Profiles), and because of the greater functionality offered by the OSI applications, such as X.400, FTAM, VT (see 2.2.5).

Nevertheless, TCP/IP and OSI will co-exist for some years to come, and the transition from one to the other will be gradual. For economic reasons, this co-existence will be on the same physical networks: the same computers, circuits, routers, and other paraphernalia which make up a network. Given this fact, the first requirement must be to ensure that communication using one protocol suite does not interfere with the other. This requirement has already been met, at least at the lowest protocol layers; there are no interoperability issues at the physical layer as OSI and TCP/IP packets are identical. At the data link layer, OSI and TCP/IP packets can be simply and readily distinguished enabling them to be passed to the correct network layer implementation. This allows TCP/IP and OSI to co-exist on LANs and on X.25 networks. The problems begin at the network layer, where a number of important areas have not yet been standardized for OSI, although work is in progress. A major problem at this level is not TCP/IP interoperability with OSI, but interoperability of OSI Connectionless Network Service with OSI Connection-Oriented Network Service (see 2.2.2).

The following sections describe some of the ways in which co-existence and gradual transition might take place.

5.1.2.1 OSI over TCP/IP

The opportunities for running OSI over TCP/IP stem from the fact that these protocol suites, both based on layered models, are very similar at their lowest layers, from the Physical layer through to the Transport layer. In fact, TCP/IP now generally uses OSI protocols at the Physical layer and the Data Link layer (X.21 and HDLC, respectively). Because of the similarity at lower levels, it has been suggested that one mechanism for transition would be to run OSI upper layers on top of TCP. This would have the effect of providing OSI transport layer services as a TCP/IP application protocol. A single network infrastructure would therefore be able to support both TCP/IP and OSI applications. This approach has been formalized in a TCP/IP document (RFC 1006) and is implemented in the ISODE application development environment for UNIX systems.

This approach does not, however, provide interoperability of OSI applications with TCP/IP applications; it merely allows OSI applications to interoperate with other OSI applications over a TCP/IP network. TCP/IP applications and OSI applications would remain entirely separate unless specific application gateways were built (Callon, Hagens, and Nitzan, 1990).

5.1.2.2 OSI gateways for existing systems

One important reason for the slow adoption of OSI, particularly among vendors of library systems, is that implementation of OSI application level protocols would require major changes in their software, often at the level of basic system design. Bibliographic applications are very complex in comparison with other business applications, and vendors are understandably reluctant to invest large sums on making radical changes to systems that currently both work well and have a high degree of acceptance.

A central problem is that, while functions are highly modular in present-day systems and data is highly integrated, both the functions and the data are often tied directly to a particular user interface. Two examples of this dependency are: 1) when the formulation of a database search cannot be separated from the screen display used to input the search terms, and 2) when data is stored in a manner that minimizes processing when it is output to the screen. In both cases, the function design and data structures are inextricably intertwined with the user interface. To convert the application from such a monolithic, centralized model to the distributed client-server model, typically required by OSI, would require that the whole design of the system be changed.

One means to overcome the problem of system redesign in the short to medium term is to leave the system unchanged and build a *gateway* to sit between the system and the OSI application. This gateway would receive OSI messages, process them internally to generate the appropriate commands for the non-OSI system, receive its replies, process them and generate OSI messages. There are, however, a number of difficulties with this approach. First of all, by adding a possibly very complex processing layer, response times may be significantly worsened and/or additional processing power may be required. A second problem is that, while it is possible to design gateways which use existing bibliographic applications as servers, it is much more difficult to implement a gateway which allows an existing system to act as a client (Callon, Hagens, and Nitzan, 1990).

5.1.3 Implementation of OSI Application Protocols

This section reviews work currently under way to implement the main ISO Application layer protocols of interest to librarians. In addition, issues raised by the current implementation work is discussed.

5.1.3.1 Z39.50 and SR

Implementation of SR origin systems and target systems need to be considered separately, since they tend to have quite different communities. While an SR target is appropriate for an information provider, an SR origin is appropriate for an information user.

SR target systems

If we look at target systems first, there are a number of implementations already nearing completion. These generally have been undertaken by the large bibliographic utilities, although

at least one university (Virginia Polytechnic Institute and State University) is intending to provide integrated access to all its information services through use of an SR target implementation.

Most of the current target implementations are designed primarily to support the exchange of bibliographic records. This is due, in part, to the fact that the standards-making process has been driven largely by the providers and large-scale users of these records, such as the Library of Congress and the Research Libraries Group. A high degree of standardization of the bibliographic exchange record has resulted, thereby minimizing problems associated with different data structures. SR targets are currently being implemented for general information retrieval systems. For example, both the University of California at Berkeley and Thinking Machines Corporation are implementing SR targets to access full text databases. Integration of SR target implementations more generally into library systems can be expected to take longer, primarily because of the incompatibility between most catalogue access systems and the SR model described earlier.

The focus on bibliographic records has meant that standards for record compositions of other kinds have not been developed. This has lead to the view in some quarters that the protocol is exclusively designed for this purpose. As a consequence, other protocols are currently under development for other kinds of database retrieval. For example, the TCP/IP community is engaged in defining a protocol to support "Campus-Wide Information" services (CWIP). Another group is defining a standard for use with CD-ROMs, "CD-ROM Read-only Data Exchange" (CD-RDx). Coordination of these activities is urgently required.

SR origin systems

SR origin systems are likely to be more widely implemented than SR target systems in the long run because they will provide the interface between the user at an information workstation and the whole set of databases to which the user has access. As outlined in 4.1.1, such an approach has an enormous appeal. The library can provide access to both local and remote information sources with a consistent interface. Origin systems, however, are likely to be more problematic to implement than target systems because they must cope with the different capabilities of disparate target systems while providing a coherent interface to the user. There are many questions surrounding origin systems that remain to be addressed. For example:

- how will an SR origin system determine what capabilities are available at any given SR target system?

- how far should the SR origin system guide the user's search strategy, with regard to both the choice of the database and the formulation of the query?

The answers to such queries depend in part on the quantity and quality of SR targets which are readily accessible. Even now, the range of organizations currently planning origin implementations is wider than those planning target systems, including both large bibliographic utilities and national libraries, as well as developers of commercial PC and workstation systems.

5.1.3.2 FTAM

Commercial implementations of FTAM are becoming available. At the beginning of 1990 there were at least 10 FTAM products, with more in the planning stage. These initial implementations, however, seem to offer rather less functionality than the full standard. American implementations, for instance, generally conform to the *NIST Stable Implementors' Agreements* which requires support for only a limited subset of the file types defined in the standard. The minimum implementation defined requires support for the transfer of only complete ASCII or binary files.

To be able to use FTAM for such purposes as exchanging bibliographic records or holdings information, implementation of a richer subset of FTAM services will be needed. Although these richer subsets are required in subsequent implementation phases, it will be some time before the full potential of FTAM for library applications is widely realized.

5.1.3.3 X.400

While there is a large and growing number of X.400 products available in the marketplace, almost all of these implementations are based on the 1984 X.400 protocols rather than the more recent 1988 version. The chief differences between the two are that the 1988 standard supports a richer set of Interpersonal Messaging (IPM) body part types, supports security features, and introduces the concept of a "Message Store" to provide a mailbox capability. A number of suppliers, however, are currently in the process of implementing the 1988 standard, and products are becoming available.

One problem with many current implementations of X.400 to provide other application software with the capability to initiate messages. The current draft profile for an ILL application, for instance, specifies that an ILL message is to form a body part of an X.400 message. It will therefore be necessary for the ILL application to make use of X.400 services to automatically create and send the appropriate messages. Similarly, at the receiving end, the X.400 application must be able to route the contents of messages to an ILL application when appropriate. To this end, a group of software developers, the *X.400 Application Program Interface Association*, are currently developing specifications for standard interfaces to X.400 applications. This has the benefit of allowing developers of other applications, such as ILL, to design their implementations independent of the X.400 applications which may be used in conjunction with their applications.

The one interface which has been standardized is for an "X.400 Gateway" intended to allow non-X.400 electronic mail systems to interoperate with X.400 systems. The specification is sufficiently generalized to provide general Message Transfer Agent services to other applications.

5.1.3.4 X.500

One impediment to the more widespread use of the *Internet* by the library community is the absence of an online directory to facilitate the retrieval of network addresses for institutions,

applications, and individuals. Currently, a user is required to rely on prior knowledge obtained in some indirect manner, such as published lists (which might be paper or electronic in form), or referrals from other knowledgeable users. In the OSI environment, the X.500 standard is intended to remedy this situation, by permitting user-friendly access to address information.

At present, there are a number of vendor implementations of the Directory in existence, but the most widespread use of the Directory is still in the context of pilot projects, the best known of these being the *Internet* "White Pages" pilot project. This project uses an implementation of X.500, called QUIPU, that operates over the *Internet* TCP/IP suite of transport protocols. It focuses on supporting name to address mapping for electronic mail purposes.

The Directory is considered by many to be critical to the successful deployment of OSI, but there are many issues that will influence its usefulness. Among these are:

- ***widespread availability*** It is important that the Directory eventually contain information about the entire networking community. Its usefulness depends on a "critical mass" of information being accessible via the Directory

- ***information content*** The base X.500 standard defines some basic object types and attribute types to be supported by the Directory, and suggests a particular Directory Schema. There has been considerable debate, however, about what information is suitable for inclusion in the Directory. Some parties advocate a restricted focus on name-to-address mapping, while others suggest that the Directory can be used as a repository for almost any kind of information

  The *Internet* White Pages pilot project has focused on providing a name-to-address service in support of electronic mail. A proposed naming architecture has been developed that identifies the basic set of object classes and attribute types to be supported. This set is a superset of the one in the base X.500 standard. Similarly, the U.S. National Institute for Standards and Technology (NIST) is sponsoring an X.500 pilot project to examine directory requirements in the U.S. government and has defined a directory schema for this purpose

  There has been considerable interest in exploring the potential use of the Directory for disseminating information. The National Library of Canada, for example, established of a pilot project to experiment with a multi-site directory based on a schema that is an extended version of the *Internet* schema. This schema includes library-specific information to support ILL and information retrieval

- ***standard incompleteness*** The 1988 version of the X.500 standard is incomplete and unusable in some important respects, including access control and replication (i.e. how directory information is distributed over more than one site). As a result, some early implementations have developed ad-hoc solutions, resulting in interoperability problems. Work is currently underway in ISO and CCITT to overcome the perceived shortcomings,

but until these are resolved and the needed changes implemented widespread availability of the Directory will be hampered

- ***support for multi-stack environments*** Since it is anticipated that both TCP/IP and OSI networking environments will coexist for some time, the Directory can play an important role in helping a system to determine what protocol stack to use when communicating with a particular destination system. The Directory can also serve to identify an appropriate gateway system if the originating and target systems do not share a common protocol stack.

  For this to be possible, it will be important that Directory services available on the *Internet* be compatible with those available in the OSI environment.

5.1.3.5 The ILL protocol

One of the factors that is likely to delay the widespread implementation of the ILL protocol in North America is the fact that most of the major bibliographic utilities, such as OCLC and RLG, already have ILL messaging services, generally using proprietary protocols and data structures. Many will be reluctant to change to the new standard when the systems they already have work reasonably well. This may be counterbalanced by strong pressures at both the national and international levels in support of standardization. The National Library of Canada, in particular, which has been a champion of the standard, is encouraging a number of vendors to produce products which implement the standard. In the international arena, the International Forum on Open Bibliographic Systems (IFOBS), which includes both national libraries and major utilities, has also been active in developing the standard.

There are several international ILL pilot projects taking place. The National Library of Canada, the British Library Document Supply Centre, and the Library of Congress are participating in the ILL Demonstration Project to demonstrate the use of the ISO ILL protocol for the exchange of ILL messages in the international arena. Another involves three European Community countries, the United Kingdom, Netherlands and France, each represented by one regional or national interlending agency.

An issue that remains to be resolved is the mode of communication required to support the operation of the ILL protocol. Manual ILL systems were almost universally based on traditional postal services and analogues such as telex. The network analogue is a store-and-forward messaging system such as X.400. The ILL protocol therefore specifically allows such systems to be used as the underlying communications service, and it is expected that many implementations will make use of this. Some bibliographic utilities that are implementing ILL, however, are committed to connection-oriented protocols for their other services and will wish to use connection-oriented communications for ILL messaging. This raises considerable problems of interoperability between ILL systems using incompatible communication modes.

Use of ILL with a store-and-forward messaging system also raises problems of integration with other library applications, such as SR, which are connection oriented. One common scenario would be that of someone using SR to search a library catalogue to locate an item and then issuing an ILL request for the item. While the two protocols can in principle be used together in a single application, the fact that they use different communication paradigms makes the task of integrating them considerably more difficult.

5.1.3.6 Virtual Terminal

The key issue surrounding the Virtual Terminal protocol is to gain its general acceptance in the face of widespread use of the TELNET and Triple-X (X.3/X.28/X.29) protocols. To achieve greater acceptance, VT-profiles have been developed that resemble the operating characteristics of TELNET and Triple-X. This change provided two migration possibilities:

- the host and server can convert from their existing protocol environment to the VT-based one without affecting existing applications since the new and old environments are functionally the same

- Virtual Terminal gateways can be established that provide a mapping from TELNET/Triple-X to the corresponding VT profile. This allows, for example, a user to login to the gateway host via the virtual terminal protocol and then login to other hosts using TELNET or Triple-X. Since the selected VT-profile is functionally the same as the other protocol, no information is lost during the mapping

  This approach has the important advantage that it allows interoperability among VT-based host or client systems and the large installed base of systems that use the other protocols, without requiring all such systems to convert.

5.2 Policy issues and other considerations

The development of information networks will require the resolution of many policy issues in addition to the technical issues discussed above. The use of research networks by libraries will affect many of the established patterns, processes, and organizations associated with information production, access, and distribution. Such changes will require the modification of existing policies and the development of new ones that relate to these domains. Some of the more serious issues to be resolved include network governance, intellectual property rights, appropriate use of the networks, convergence of library and computing centre operations, and issues relating to end-user policies and staffing.

5.2.1 Governance

As research networks begin to evolve into information networks, providing access to libraries and other information resources, and accruing larger and more diverse user communities, a major issue that confronts developers is how to provide overall governance for this expanded form of network. *Governance* has been defined as the processes through which an organization "confirms its goals, establishes policy, selects its leaders, and insures the necessary resources to carry out its purposes" (Mathews, 1979). Governance is a primary concern because the creation and maintenance of a national information network will require the participation and coordination of many diverse groups. The two most germane to this discussion are individual research networks that comprise large internetworks, and library organizations devoted to resource sharing, such as library networks and consortia. If libraries are to use research networks to support bibliographic information provision, document supply, and technical services, then some form of overall governance must be established that can adequately represent these two large groups.

Establishing overall governance for an information network will be a considerable challenge because of a number of obstacles:

- the national research networks that will form the basis of information networks have problems of their own in establishing governance
- research networks and library networks must be reconciled in some way in order to map current resource sharing arrangements onto the framework of research networks
- there are numerous of cooperative organizations that libraries have created to facilitate resource sharing that must taken into consideration
- all parties concerned have differing sets of priorities, needs, interests, goals, policies, funding sources, geographic locations, and existing forms of governance.

The severity of these obstacles, however, is not the same in every global region. In Europe, for example, national research networks and library networks are already more integrated, facilitating the effort to develop an overall governance structure. In North America, on the other hand, the obstacles are more severe because there are a large number of distinct research networks and library networks to be represented within one overall governing body. These two regions will be discussed separately below.

5.2.1.1 Governance in the U.S.

Research networks

The first major obstacle to overcome in establishing overall governance in the U.S. is that the *Internet* does not have a centralized body to direct, manage, and create policy; there currently is no consolidated form of governance overseeing its development. Instead, the *Internet* is a patchwork of separate networks that have been interconnected in largely a haphazard and *ad hoc*

manner (Lynch, 1990). Component networks, originally developed for different purposes and to serve specific user groups, have been interconnected without regard for the governance of the overall structure.

For most of the individual networks making up the *Internet*, governance is provided in the form of a board of directors comprised of representatives from the universities, colleges, and research centres that make up their membership; each is responsible for the direction and policy decisions for their portion of the *Internet* (Quarterman, 1990). As a consequence, the policies, usage guidelines, and centres of responsibility controlling the component networks vary widely across the *Internet*. This diffuse form of governance creates serious problems in controlling and directing the internetwork as a whole (Mandelbaum, 1990). Such problems are, for example, no feedback mechanism to respond to end-to-end network problems (those that cut across the boundaries of individual networks), no authoritative body to standardize and enforce policies or guidelines across the entire *Internet*, and no body to respond to the needs of users. This situation has prompted one commentator to remark that the current *Internet* is, as a whole, "undirected and largely unmanaged" (Mandelbaum, 1990:11).

Restructuring of management organizations has been attempted to better manage the rapidly growing *Internet*. One approach involves the creation of a not-for-profit organization, Advanced Network and Services Inc. (ANS) to manage and operate the NSFNET backbone, one of the two main backbone networks of the *Internet*. The formation of ANS by Merit, IBM, and MCI, is designed to more effectively manage the backbone; it will supposedly make the networking infrastructure "broader and more productive" and provide increased access to "information technology resources, such as national database centres, medical imaging, satellite data, and supercomputer centres" (News Release received over BITNET, 1990).

A second approach has been a controversial "grass-roots" movement to commercialize the *Internet* proposed by a group of users who were dissatisfied with the management structure and costs of the present internetwork (Jackson, 1989). "Commercialization" in this case does not mean commercial access to the *Internet*, but its management by an independent, for-profit organization. Many believe that turning the *Internet* into a for-profit enterprise, allowing users to buy the services they require, is the most promising way to improve the overall governance of the internetwork (Mandelbaum, 1990). Others, however, believe that making the *Internet* a for-profit network will destroy the casual, open, university-like atmosphere that promotes discussion and research (see Jackson, 1989). The first step in commercializing the *Internet*, however, has already taken place. Early in 1990 Performance Systems International Inc. (PSI), a provider of value-added network services, signed a five year contract with NYSERNet Inc. to take over the management of that portion of the *Internet* (Jackson, 1990). While it is too soon to judge the effects of such a move, the fact remains that a new consolidated form of governance will be required to turn the existing *Internet* into the high-speed, high-capacity information network that has been proposed.

Yet another approach is to use a group of technologies called *policy-based routing*. These technologies are under development to help better deal with the policy and management problems

of carrying data across a complex internetwork by embedding use restrictions (see 5.2.3) in the workings of the network itself (Lynch, 1990). For example, traffic may pass through a particular route only if has been previously authorized to do so.

Library networks/consortia

Solving the management and policy problems of the research networks alone, however, may not be sufficient to develop an information network. It may also not be sufficient to enable research networks to support library functions. A full solution may require that library cooperative agreements be taken into consideration in any overall governance structure. That is, if libraries are to use research networks to support bibliographic information provision, document supply, and other services, then pre-existing cooperative agreements should be extended to the research networks.

The cooperative arrangements in which libraries engage are the result of a long-standing tradition of resource sharing. One source states that "cooperation and reciprocal functions of various kinds have become commonplace among U.S. libraries in the last 75 years" (Martin, 1986:11). This resource sharing was initially carried out by several hundred traditional, non-automated library consortia until the 1970's, when they were largely replaced by automated library networks. These networks are the various bibliographic, regional, and local library networks that support cooperative activities such as reciprocal borrowing privileges, coordinated acquisition policies, preservation programs, shared cataloguing, (Martin, 1986), document delivery, contract cataloguing, union lists of serials, access to online databases, email, and collection analysis (Shaughnessy, 1987).

Given this long-standing tradition of interlibrary cooperation, it may be advantageous to preserve aspects of existing cooperative arrangements and the organizations that support them; pre-established resource sharing arrangements should be mapped, where possible, to the research network infrastructure. This will require that a model of overall information network governance include libraries and their cooperative organizations.

Coordinating research networks and library resource sharing arrangements, however, will not be a trivial exercise. There is a tangled web of library networks, consortia, and other cooperative organizations, with which to deal. The array of terms used to describe cooperative organizations, such as *library networks*, *library consortia*, *library associations*, *library systems*, and *library cooperatives*, attests to their great variety and number (Walters, 1987). The number of cooperative organizations that meet the simple criteria of "dependency on computation, telecommunications, and machine readable data", for example, is well over 300 (Walters, 1987). These library networks vary on a number of dimensions: the size of their membership, the scope of the geographic coverage, financing, and in their association with like libraries. (Shaughnessy, 1987).

Walters (1987), in his study of local library consortia has identified fully five different categories of library networks:

- ***unit cost networks*** These are mainly bibliographic utilities such as OCLC, RLIN and UTLAS. They were created because of the need to reduce the "unit costs" of functions such as cataloguing, acquisitions, serials control, and interlibrary loan

- ***multi-state regional auxiliary enterprise networks*** These are the large regional networks that provide computerized services to members by contracting collectively with OCLC. Examples are Amigos, Capital Consortium (CAPCON), the Illinois Library Network (ILLINET), and the Southeastern Library and Information Network (SOLINET) (Martin, 1986). They are considered to be "important players" in establishing policies for national cooperative activities

- ***authority-sanctioned networks*** These networks are similar to the two listed above, except that they are established and funded by governmental authority

- ***discipline and type-of-library networks*** Membership in this type of network usually stems from a relationship to a specific discipline, like Law or Business, or to a type of institution such as a university or corporation

- ***local consortia or proximity networks*** The organizational principle for this type of network is the proximity among members. Within this category, individual networks can vary as to size, organization, form of governance, and programming activity

The diverse set of interests represented within both the research network community and the library network community will make the establishment of governance very difficult. One of the primary issues involved with creating the proposed National Research and Education Network (NREN) in the U.S., in fact, is how it will be governed (see Chapter 6). It may be that the resource sharing activities supported by only a selected number of library networks may be manageably incorporated. Limits will probably be set, at least initially, by the physical capacity of the research network and by the fact that only certain types of libraries, specifically those at academic, governmental, and research institutions, will have access to the research network.

5.2.1.2 Governance in Europe and the U.K.

Research networks

The development of centralized governance for European and British national research networks is farther advanced than it is for the American *Internet*. One reason for this difference, no doubt, is that the geographic scope of national research networks in Europe, and JANET in Britain, is much less extensive than that covered by the *Internet*. The size of a typical European national network would probably be comparable to regional network of the *Internet*. Another reason is that Europe and the U.K have had central funding that is "more secure and directed" than their U.S. counterparts (Dempsey, 1990:13).

The literature on the topic of the governance of European research networks taken as a whole is extremely scant. However, the initiation of a pilot project for a pan-European network called IXI (Pilot International X.25 Infrastructure Backbone Service), indicates that there are efforts to interconnect and coordinate the activities of European research networks (IES News, 1990a). The IXI pilot project is an effort by the Commission of the European Communities (CEC) and PTT Telecom of the Netherlands to connect all European and British research networks to an X.25 backbone. It is the first major activity of the EUREKA COSINE project (Cooperation for Open Systems Interconnection in Europe), a consortium of European countries created to encourage implementation of ISO-OSI networking. It is governed jointly by the COSINE Policy Group, representing CEC national governments, and RARE (*Réseaux Associés pour la Recherche Européenne*), representing the actual and potential network user community (Quarterman, 1990). PTT Telecom will be responsible for creating, operating, and managing the internetwork in cooperation with other European PTTs (IES News, 1990a). While governance of the IXI internetwork will be provided by COSINE for the short term, there are uncertainties about governance and funding beyond February 1991 when the Pilot Service Phase expires. The COSINE Policy Group intends to ensure, however, that the facilities remain available beyond this period. The manner in which this will be accomplished has yet to be determined.

Library networks

The establishment of governance for a European information network should also include the representation of European bibliographic networks. The situation in Europe is similar to that in the U.S. in that there is a wide variance in the size, scope, organization, and services offered by European library networks (Dempsey, 1990). There are, however, some important differences. The first is that there are over 30 bibliographic networks in 14 European countries. This is about one tenth of the number operating in the U.S. It will doubtless be an easier task to represent the interests of 30 library networks as opposed to 300 in a model of overall research network governance. The second difference is that there is *already* a close association between research network and library networks in Europe. Pica, for example has recently signed an agreement to make Pica's library network part of the SURFnet infrastructure (Buitelaar, 1990). Similar examples include the close associations between BIBSYS and UNINETT in Norway, LINNEAS and FUNET in Finland, CURL and JANET in the U.K, as well as the plans for the SIBIL library network to use SWITCH in Switzerland (Dempsey, 1990).

Establishing an appropriate form of governance for research networks is crucial if they are to provide the advanced services and the wide variety of information resources that have been proposed. Because libraries in both the U.S. and Europe have established elaborate cooperative agreements with other libraries to share resources, and that these resources will be a valuable part of the information networks that will eventually develop, then participation by libraries should be reflected in the governance structure. At this point in time, however, the final form of overall. governance structures has not been determined.

5.2.2 Intellectual property rights

A major stumbling block associated with providing information resources over research networks, and indeed with all new information technologies, is the issue of intellectual property rights. This issue concerns the fact that the existing body of laws governing intellectual property rights, such as patents, trade secrets, unfair competition, and copyright, have not kept pace with rapid advances in information technology (Miller and Blumenthal, 1986). Basic assumptions that were true in a paper-based environment are no longer true in an electronic environment. Of particular interest to libraries are those intellectual property laws protecting original works of authorship, known as *copyright*. The inadequacy of this body of laws is reflected in the fact that litigations within the information industry have risen dramatically from approximately 600 cases in 1981 to over 5000 in 1985 (Goldstein, 1986).

Copyright

Copyright is based on the premise that individuals should be compensated for their labour; that is, authors, artists, and other creators of intellectual "property", should be protected from the unauthorized copying of their work (Miller and Blumenthal, 1986). Because ideas are intangible, copyright protection has traditionally been applied to the *tangible* expression of ideas in permanently fixed forms, such as film, sound recordings, and print materials, rather than the ideas or information themselves. The notion of tangibility has formed the basis for compensating authors for particular expressions of ideas because the relative value of the ideas could be determined by the number of copies made of the fixed form; the greater the value of the information contained in a book, the more copies of it are made and sold. Because of the making and copying of fixed forms has traditionally been a specialized and time-consuming process, the number of copies of a work made was easy to monitor and control.

This assumption is no longer true, however, with the advent of modern technologies. For example, photocopiers were one of the first technologies to put pressure on the copyright system because multiple copies could be made quickly and effortlessly. This diminished the control over the reproduction of a work, hence the basis for compensation. The balance of the scale tipped away from the information producer toward the information user who could now make as many copies of a work as he or she wished. To redress the balance, the concept of *fair use* was adopted to attempt to control unbridled copying. The fair use doctrine permits the reproduction of selected parts of printed documents for personal use (Cartwright, *et al*, 1989). Fair use was an attempt to provide adequate compensation to information producers, while allowing the free flow of information to users.

In an electronic environment, the problem of monitoring and controlling the distribution of copies becomes even more acute. Electronic information does not have to be printed, but exists in a form that can be easily copied and "transferred from user to user, or machine to machine... totally beyond the control, or the knowledge, of any third party" (Miller and Blumenthal, 1986:229). Electronic media can be so easily copied, in fact, that it has been labeled "dynamic media" (see Avram, 1989). Thus, controlling copies, the basis of compensation, breaks down completely with such media. The balance has been tilted to such an extent toward the information user that

producers are seeking alternate means for compensation for information shared by users (Avram, 1989). These alternate methods include pricing schedules that attempt to anticipate the rate of copying, government imposition of compulsory royalties, and taxation of blank media.

Copyright in a network environment

The problem of copyright and the control of multiple copies reaches dramatic proportions in a network environment. The dynamism of information in a network environment, and the ability of multiple users to simultaneously access and make their own copies of information, makes the application of copyright law extremely difficult (Avram, 1989). A few of the more problematic areas upon which current copyright laws may have an adverse impact are discussed below.

Network-accessible databases Under copyright law, the unauthorized copying of information or its arrangement is an infringement of copyright. Online databases, however, are seldom copied in their entirety; only bits and pieces of information considered to be useful are extracted at one time. These can then be reformatted into an almost infinite number of new combinations or arrangements and may thus constitute a new copyrightable work (Miller and Blumenthal, 1986). If users create their own personal databases, or sell a particular version they have created (e.g on a specialized topic extracted from multiple databases), it may be unclear under current copyright doctrines whether copyright has been violated. Producers of the original databases will certainly consider it to be so. In a network environment, with many information sources accessible to a workstation capable of sophisticated manipulation of data, the creation of these secondary databases will become a casual endeavor. Moreover, when the information on networks become truly distributed, there may be no way to trace all the possible sources of information used to create a secondary database.

Electronic document delivery The impact of current copyright law on electronic document delivery is even more profound. In current photocopy-based document delivery systems, there is already considerable concern over the copyright issue; libraries must keep detailed records of every photocopy requested to avoid substituting for subscription or purchase and thus violating copyright. In electronic document delivery, the whole concept of delivering documents electronically is based on making a copy of a work every time it is needed. In a research network environment, where literally millions of copies of a work may be sent to millions of users, developing strategies for fair compensation becomes crucial. Current copyright laws are clearly inadequate to deal with such a situation. One proposed solution to this problem is to embed compensation mechanisms within electronic document delivery systems (Kost in Avram, 1989). In such a system, debits and credits could be applied to the accounts of information producers, distributors, and consumers automatically based on the use and rates of copying of a particular piece of information.

Imaging An issue related to electronic document delivery is the problem faced when making digitized copies of paper documents. If paper-based materials to be digitized are copyrighted, then the act of making exact digital copies violates that copyright. While scanning a copy for personal use may fall under the fair use doctrine, wholesale conversion, such as would be done when creating an image database, does not. The conversion process is considered to be a

"distinct category of reproduction that would require a separate authorization from the relevant copyright source" (Lyons, 1990:9). Since one major source of documents for electronic document delivery be will be digitized images, at least for materials originally in paper form, copyright issues will need to be considered at two stages: 1) during the creation of the initial digital image and, 2) when copies of the image are made available for distribution over research networks.

Cataloguing Copyright entered the world of shared cataloguing in 1984 when the Library of Congress registered a claim by OCLC to copyright its union catalogue (Lowry, 1987). The issue of copyright of this bibliographic database is of "profound importance to the library community" because "literally thousands of libraries participate in OCLC services" (Lowry, 1987:81). It also affects those libraries who do not deal specifically with OCLC. OCLC's "Principles and Guidelines" state that, while libraries may transfer records to other member libraries without restriction, they are not allowed to transfer records to "member networks, local consortia, or state library agencies" unless they have made a specific agreement with OCLC (Lowry, 1987:81). This has the effect of prohibiting libraries from "entering into arrangements for resource sharing, (e.g. ILL, retrospective conversion, and union catalogs, etc.)" with other organizations without written permission from OCLC (Lowry, 1987:81). This copyright restriction strikes at the heart of many of the library uses of research networks: that is, ILL, shared cataloguing and other resource sharing activities that research networks have the potential to enhance. If the records are originally derived from OCLC, then such resource sharing may not easily take place.

Authorship The ability to exchange processable documents among authors over research networks will enhance and encourage collaboration. While this may be beneficial to scholarship, it will dilute the concept of "first author"; that is, when a document can be passed repeatedly from author to author in processable form, individual property, and thus economic compensation, will be difficult to determine (Cartwright, *et al*, 1989). Research networks will also enable new forms of joint works to arise. These are, for example, multimedia documents made possible by ODA, a document model specifically designed to facilitate interchange and remote access to documents. ODA documents can contain text and static images, and will eventually contain moving images, voice, and sound, each of which may have different authors and copyright restrictions. Moreover, these documents may be combined with hypertext, embedding multimedia documents within multimedia documents, resulting in an inextricable tangle of distinct copyright restrictions. It will be extremely difficult to apply current ideas of authorship to such unprecedented forms of documents.

International concerns

While domestic copyright laws governing electronic information may be inadequate, indeterminate, and confusing, the problem is even more extreme in an international context. This stems from the fact that it is difficult to synthesize different countries' copyright laws into a unified international standard (Miller and Blumenthal, 1986). Even though most countries have laws covering traditional forms of expression, they vary as to the protection they give new forms, such as electronic information. Furthermore, tension is growing between developed and

developing countries in that the former wish to protect the information that is becoming an increasingly lucrative commodity, while the latter seeks greater access to the information it needs to progress economically and socially.

5.2.3 Appropriate use

As research networks become more ubiquitous and offer a wider range of services, a severe issue that faces administrators is the "appropriate use" of these networks. Traditionally, most academic networks have usage guidelines that restrict the types of activities in which users can engage. These include restrictions on subject matter, language, or certain resource-consuming activities like the transfer of large files. Common to almost all guidelines are restrictions on for-profit usage by competitive enterprises, otherwise known as *commercial use*. BITNET, for example, allows a broad range of users and uses, but restricts transfer of information to that which is noncommercial in nature or supports its member institutions' academic missions (Fuchs, 1988). It does not allow job solicitations, transmission of proprietary software products, or any other type of commercial usage. NSFNET's guidelines, as well, state that the network is to be used for the support of research and instruction at non-profit organizations, but the rules do not explicitly forbid its use by industrial sites if they support the mission of NSFNET or its members (Fuchs, 1988). The grounds for prohibiting commercial use of research networks are generally that research networks are publicly-funded, and the use of them to either advertise a product or service, or to enhance an existing commercial service, is a misuse of those funds and represents a form of unfair competition (Fuchs, 1988).

There are two central problems, however, with the restriction of commercial use of research networks. The first concerns the definition of commercial use within the guidelines themselves, and the problems of their application and enforcement on the network. The second problem is more fundamental pertaining to whether or not research networks can prohibit all commercial activity and still provide the advanced services, such as multiple database access and electronic document delivery, that are proposed for future networks.

Definition, application, and enforcement

"Commercial usage" is a very broad term that can refer to many different activities in different contexts. While advertising and junk email are clearly commercial, other types of activities may be less clearly so. For example, many research networks allow private research companies and industrial sites to be members so long as they are working on projects associated with the universities. If, however, the companies eventually make a profit from the work that originally supported academic research, it may constitute commercial activity. Statements such as "to advance scientific collaboration on a national scale" or "support an academic mission" are open and vague enough to allow broad interpretation. If commercial usage is too broadly defined within usage guidelines, it could restrict many valuable activities. A narrow definition could conversely allow inappropriate use to occur. It many situations, it is up to network committee members to decide what is proper use on a case by case basis (Fuchs, 1988).

A second problem is inconsistency in usage guidelines across large internetworks. Large research networks are actually interlinkages of smaller, independent networks, each with its own set of usage guidelines. What may be allowed on one network may not be on another. Such diversity makes even application of rules against commercial use difficult. This would be especially true in an international context where cultural and political differences also affect guidelines. Groups that have formed to coordinate activities among research networks in several topic areas, such as the Federation of American Research Networks (FARNET), have targeted usage guidelines as an area for coordination.

Current methods of enforcement for guidelines on research networks are often no more than peer pressure or self-policing. Violations can result in chastising email messages (Quarterman, 1990), or removal of an institution from the network (Fuchs, 1988). As traffic increases between a multitude of parties, organizations, and machines, this method will quickly become inadequate. Compounding this problem is the fact that there are no apparent boundaries between networks; traffic travelling from one point to another may go through several distinct networks making enforcement of usage guidelines more difficult.

Commercial usage and advanced services

The issue of commercial use has the potential to become a major obstacle as research networks begin to offer a wider range of services. In the past, services offered by research networks typically were communication between colleagues, supercomputer access, and file transfer. While private firms also may have used the networks, explicit, for-profit activities would have been prohibited. However, many of the services that users are currently demanding will violate commercial usage restrictions. For example, when BITNET began establishing nodes on the west coast of the U.S. where DIALOG hosts resided, the information vendor wanted to connect its hosts to the network (Fuchs, 1988). They proposed that a connection to BITNET would allow users to receive results of their searches directly through email, rather than waiting for off-line prints to be sent by regular post. DIALOG realized that BITNET could not support interactive activities. Rather, a user would connect to DIALOG through a PDN and conduct a search as usual, but the "print" command would automatically forward the results through BITNET. The proposed service would have meant greater convenience for DIALOG and BITNET users. Individuals representing BITNET at the time, however, considered such use of the network to be inappropriate because DIALOG's service would have been enhanced at the expense of BITNET.

DIALOG's proposal to BITNET was turned down because it was considered to be too commercial, yet, these are exactly the kinds of services that are proposed for future research networks. Services such as access to multiple databases, or the production and delivery of electronic documents, will necessarily involve commercial enterprises. It will be the database vendors, publishers, bibliographic utilities, and document suppliers who will provide the raw materials for the services users demand. It may become increasingly difficult, and even ***undesirable***, to restrict commercial usage in this light. The areas of greatest concern to libraries which strict commercial usage guidelines may affect are access to commercial information services, access to bibliographic utilities, proposed document delivery systems, and acquisitions systems. These areas are discussed below.

Access to commercial information services If libraries are to provide their users with access to multiple information resources through the network, then commercial vendors must be dealt with at some point. Two major methods through which multiple database access can be provided over the networks are: 1) mounting databases at a particular site and then licensing them for access in a pre-determined network area, and 2) allowing the hosts of information vendors to become database servers on the network. Both methods raise concerns about commercial usage as network access would enhance the vendors' services as well as increase their market size.

The U.K.'s research network, JANET, for example, currently licenses four commercial databases from the Institute for Scientific Information (ISI) for which they pay an annual flat rate fee allowing unlimited end-user searching (Information World Review, 1991). It is clear that the number of individuals searching ISI databases is vastly increased through their availability over JANET than would be otherwise possible. While this is a boon to end-users, the question arises of whether this vastly increased market size for the vendor is contrary to restrictions on commercial use as it is currently defined..

A problem related to vendor hosts acting as database servers on the network is that of determining fair pricing structures. As Fuchs (1988) points out, use of networks by commercial enterprises creates a potential situation in which the users (or their organizations) can be charged twice for their use of a research network. Once through levies the organization pays to create and maintain their part of the network, and a second time by paying fees to the vendors that have embedded within them the vendor's own network fee.

The conflict between commercial use and information provision, at least, may now be academic. The Research Libraries Group (RLG) and OCLC have recently begun to offer access to their information services through the *Internet* on a fee-basis (see 4.1.1.1).

Access to bibliographic utilities The issue of commercial use is less clear with regard to the use of research networks by bibliographic utilities to provide their technical services. While bibliographic utilities may not want to use research networks to support their technical services because of heavy investments in their own networks, this form of access is perhaps more overtly commercial and therefore more likely to violate the guidelines.

Electronic document delivery Questions about appropriate usage also arise in the use of research networks for electronic document delivery. The first area of uncertainty is in the creation and exchange of documents. There is some concern that the exchange of a document-in-progress between an author and an editor (or two authors) over a research network which will subsequently be published by a for-profit organization constitutes commercial use. A example would be the use of a research network for the production of a textbook. A textbook is both academically-oriented and benefits students and education, seemingly satisfying current usage guidelines. However, both the publisher and the authors stand to make a profit from its

publication. Commercial use restrictions broadly applied could potentially prohibit this type of activity.

A second area that raises questions about commercial activity is the distribution of electronic documents from suppliers to users for a fee. The fact that authors and publishers may be able to exact royalty payments for every copy of a document sent over a research network makes every document sent in this way a commercial transaction. To continue the above example, the text produced through collaboration on the network could be sent to students in electronic form through the network. The students, of course, would have to pay for their copy. Because the method of distribution for the text is the research network, the publisher and authors are benefiting from the use of the network.

Acquisitions The process of acquisitions necessitates communication between a library and a commercial enterprise, such as a book seller, jobber, or publishing house. Using standards such as X.400 and EDIFACT, book orders and payments can be exchanged using the research networks as the telecommunications channel. While the use of the networks may be more convenient and cost-effective for the library, all activity between it and the suppliers is unquestionably commercial. Publishers, jobbers, and books sellers would be benefiting at the expense of the research network, representing an unfair form of competition.

It is clear that to provide the types of advanced services that are proposed for research networks, such as distributed information retrieval and electronic document delivery, commercial enterprises will have to participate at some level. It may be that restrictions on commercial use will have to be eliminated altogether and replaced by some form of licensing or fee structure to ensure fair use by all parties. In the end, maintaining research networks as a "free" service may severely limit the types of services that can be offered, if it also means keeping commercial use restrictions. Some view the shift away from the "free" services as inevitable. As Fuchs points out, "there is eventually going to have to be some way of accounting for and probably charging for use (of the research networks)" (1988:48). This may not necessarily mean individual users will have to pay, but certainly the institutions to which they belong.

5.2.4 Convergence of library and computer centre operations

The participation of both libraries and computer centres in the movement toward supplying information electronically to end-users has sparked a considerable debate as to the relative roles of these organizational entities. Originally, the library was a repository of paper-based materials that were acquired, organized, and lent to users (Neff, 1985). The computer centre, on the other hand, was responsible for machines that could input, manipulate, store, retrieve, and display information. While libraries emphasized organization and dissemination, computer centres emphasized processing (Post and Sessions, 1986). As library automation has proceeded, however, the traditional functions of library and computer centres have become increasingly interdependent. The rapid development of information technologies, especially the use of workstations to access and handle information, have blended these original functions (Olsen, 1990).

Areas of convergence

The areas of convergence between libraries and computer centres are numerous. A few found in the literature are (Woodsworth, 1988; Post and Sessions, 1986; Molholt, 1985; Olsen, 1990; Neff, 1985):

- an increasing emphasis on the storage of information in electronic form. Libraries are becoming more dependent upon electronic access to information which involves computers that can store and manipulate data
- an increasing use of workstations as the doorways to both information and computing resources
- the use of the campus-wide network to support access to the online catalogue and other information resources
- a dependence upon the computer centre for hardware, software, and network access and maintenance
- the support of data archives and other databases by computer centres
- a convergence of skills, such as database construction and local tape-mounting of leased databases
- the need to select and maintain sophisticated software.

Administration

The blurring of the roles of library and computing centres has been regarded by many as an indication that a rearrangement of organizational structure is needed (Woodsworth, 1988; Kaufman, 1990; Post and Sessions, 1986). Among the approaches to reorganization is to create a new position that subsumes both the library and computer centres. This position is that of the Chief Information Officer, or CIO, who has the primary responsibility of "coordinating the implementation and use of information technologies". This includes the computing and telecommunications that support academic computing, library and information systems, administrative systems, audio-visual operations, and office automation (Woodsworth, 1988:38). A variation on this approach is to make only library technical services and automation units part of the CIO's responsibility, leaving traditional library services and collections separate. The organizational integration of library and computer centres has actually taken place at some universities, such as Columbia University which has created the Scholarly Information Center out of these departments (Kaufman, 1990).

Other authors suggest that there is no need for a merger of these departments (Olsen, 1990). Their view is that the fundamental missions of the library and computer departments are still essentially the same, while only certain procedures have changed; the library's mission remains

"to select, store, organize, provide access to, and train patrons in the use of information" and the computer centre's mission continues to be "to provide the computing and networking infrastructure to support all aspects of academic life" (Olsen, 1990:238). Since this infrastructure equally supports the automated processes underlying library activities, instructional programs, research needs, and administrative processes, there is no reason to merge the library and computer centre and not the other organizations.

A need remains, nevertheless, for some form of closer association between the library and computer centre, no matter how it is achieved. As libraries continue to expand the number and variety of electronic information resources that they support, questions of responsibility will have to be resolved. Examples of problematic areas are:

- negotiation with vendors for search and retrieval systems
- installation and upgrade of systems
- database design and loading
- design and creation of front-end software
- consulting and training of users.

Closer association, in any form, will be needed to share skills and expertise, and to ensure consistent information technology policies, strategies, and acquisition of equipment. Such cooperation will lead to greater efficiencies in applying the technologies common to their missions (Woodsworth, 1988). This is especially true as libraries become network resources and must work with the computing centres that ultimately create, manage, and maintain network access.

5.2.5 Impact on user policies

As libraries move into the more open environment created by linking their systems to research networks, there will be a need to create a number of new polices and programs to address library users. Current policies have developed within a relatively closed library environment that is characterized by isolated library systems, fixed numbers of terminals, limited hours of library operation, and the availability of identification techniques, such as requesting student cards, to control access to certain services. This closed environment has given libraries a large measure of control over access to library systems and information resources, as well as the form and content of end-user training programs needed to teach proper usage. Within the wider, more open research network environment, however, such control will no longer be possible. The two most apparent changes to the environment will be that:

- the size and diversity of a specific library's user group will increase substantially, well beyond the library's traditional local constituency

- the range of devices available for use by local users will expand to include the set of all those connected to the network.

This larger and more diverse user group will necessitate new policies to control the increased demand upon finite computing resources as well as to limited access to information resources for which authorization is needed (Lynch, 1989b). Further, training programs will have to be extended to include instruction for the new set of systems that will be available. Both changes will be exacerbated by the fact that the range of information resources available on library systems have begun to expand beyond the local catalogue to high-demand resources, such as locally mounted databases (see Potter, 1989). The specific areas for which user policies must be developed to deal with these changes include those for management of system capacity, authorization and access, and end-user training.

Management of system capacity

Any library system has a limited capacity for the number of users that it can support at any one time. Variables affecting this capacity include the size of the processor, the size of the database, the number of terminals connected, and the expected number of transactions per terminal. An appropriately designed system that takes these factors into account can serve its community with acceptable response times and avoid overload. Within a closed library environment, the management of system capacity is facilitated by the fact that there is a finite number of terminals connected to the system and that periods of heavy use are determined by in-library activity within set hours of operation.

Once a library system is connected to a research network, however, a new set of variables suddenly affect the system making the management of system capacity more complex. Within this wider environment, determining the potential number of terminals connected to a system, the transactions per terminal, and peak load will be extremely difficult. Examples of variables affecting the performance of a system are listed below (Lynch, 1989b):

- there will be a wide variance in the number of potential users attempting to use a given system at one time. A library's potential user community in the research network environment is the total population of network users. This means that the number of terminals that may be attempting to connect to a system will be "effectively infinite" (Lynch, 1989b:13)

- network users with terminals in their workplaces will want to determine their own hours of operation, demanding access 24 hours a day, seven days a week

- because network users will come from different time zones the notion of what is a "peak period" will change

- users with windows-based interfaces will be able to maintain a connection to a library system indefinitely. This will enable them to initiate a search at any time by simply clicking on the window, but will cause many idle connections

- systems may be overloaded by script files designed to perform automatic information retrieval in the absence of the end-user.

While the solutions to these problems will ultimately be technical in nature, it is important for libraries to "establish policy objectives before considering technical solutions" (Lynch, 1989b:15). Such policies will ultimately affect users' access to the systems. Various alternatives for controlling overload include:

- ***limit the number of concurrent users*** This method involves blocking user access to a system once some level of a criterion for system performance has been reached. For example, the system could refuse further sign-ons once response time dropped below some pre-set level. Inherent in this method is a "first come first served" policy since it does not differentiate between types of users or tasks; in-library users would receive equal treatment as remote users and specialized databases would be as available as the online catalogue. While this method may be egalitarian, it may not fit with a library's current user policies, such as giving preference to faculty over students, or to institutional users over non-institutional users (Lynch, 1989b)

- ***prioritize access*** An approach that is more sensitive to user policies is to limit the number of concurrent users based on some scheme for prioritized access. This method involves segmenting the user population into distinct groups and then determining a rank order for them based on policy. The groups could be, for example, staff, faculty, students, as well as users whose searches originate from a given location, such as from within the library, or from campus-wide, regional, or national networks. In this way, when system performance drops below a certain level, specific groups could be denied access beginning with those lowest on the priority list. For example, priority could be given to staff members over faculty and students, or to local campus-wide network users over regional or national network users.

  Setting policy for user priorities will undoubtedly be a "difficult and politically sensitive process" (Lynch, 1989b:14). There will always be parties that will disagree with a certain policy. Nevertheless, libraries have had to set such contentious policies before, such as differential book loan periods for undergraduates, graduates, and faculty. The sudden expansion in size and diversity of a library's user group, and the fact that not all potential users can be served simultaneously, clearly necessitates the partitioning of the population into groups of varying priorities

- ***limit access to certain databases*** Another alternative is to selectively limit access to certain databases. The trend toward the inclusion of non-catalogue databases in library systems will result in a large number of databases that must be supported by a system. During peak periods, it may not be possible to provide access to all databases at once. Use statistics can determine the relative popularity of these databases, allowing those least used to be limited first when response times suffer. It is also possible to use a nested approach, combining prioritized user access with prioritized databases access. For example, access to particular non-catalogue databases could be denied first to users outside a network region

followed by those in a more local region. In this way, users that are closely associated with the hosting institution can access most of the databases most of the time.

Authentication and access control

Because a library's total community will be much larger and more diverse within a network environment, security, in particular authentication and access control, become crucial issues. *Authentication* involves determining the identity of a user, while *access control* involves providing or denying access a resource based on that identity (Lynch, 1989b). Authentication and access control form the foundation of the protection and management of information and system resources. In a library setting, the main functions they serve are:

- ***to protect staff functions.*** If a library system supports both public access and technical service functions, it may be vulnerable to attack or mischief from unauthorized personnel. In a closed computing environment, libraries have not had to be concerned with security since restricted functions, such as circulation, acquisitions and cataloguing, have been only available from physically separated terminals possessing simple, but adequate, security controls. In the research network environment, however, the restricted functions are not physically separate, and simple access control mechanisms are inadequate

- ***to facilitate management of system resources.*** Management of system resources requires restricting access to users during busy periods. If methods used are based on restricting access to certain groups, then authentication provides a way for the library to know into which group a given user falls

- ***to allow only authorized users to access information resources.*** Similar to management of system resources, authentication provides a method by which a library can discover whether a given user is authorized to use a particular database. This is especially important when licensing agreements restrict access to a specific subset of the network population (Lynch, 1989b).

Methods of authentication and access control Libraries must choose among various methods of authentication and access control to determine which are appropriate for their particular institutions. The main method used to identify users and control access is through assigning user IDs and passwords. This method requires all individuals who wish to access a system to register with an appropriate administrative body to acquire an account. The information obtained when registering can be used to determine a user's institutional status or point of origin, and thus can serve as the basis for computer and information resource management as discussed above.

The use of IDs in a network environment, however, brings with it a number of disadvantages. The first concerns the cost of creating and managing a database of user IDs (Lynch, 1989b). Network populations are very large, and therefore the set of potential users that may wish to search a library's database may also be very large. This problem is compounded by the fact that if a user accesses a database only once, or as infrequently as once a year, user IDs still have to be created and maintained. Keeping track of all possible IDs will require large system and

administrative overheads that may be beyond the resources of some individual libraries. Furthermore, these large databases may be repeated from institution to institution with a large degree of overlap, causing redundancy and wasted space.

A second disadvantage comes from the users' point of view. If user IDs and passwords are required for using the catalogue or other databases at every institution on a research network, then a user must go through great pains to register, update, and keep track of every ID and password required. Because of this fact, users may only keep a handful of connections active, thereby limiting their access to information. Lynch suggests that most research network users will not bother to get user IDs, thus "access from other institutions to the library catalog (is) effectively eliminated" (1989b:15). This runs counter to national policies of a commitment among libraries to share resources.

There is a clear need for authentication of users so that system management and access to information resources can be controlled, yet, there is also a need to make access to these resources as open and equitable as possible, as well as to avoid maintaining large, redundant ID databases. The solution requires a balance between uncontrolled access to resources and exhaustively recording user IDs. Lynch (1989b) offers a solution that involves examining the address of the incoming call to determine its point of origin thereby determining eligibility for restricted database use. For example, calls originating from within a campus-wide network may be assumed to be covered by a campus licensing agreement, while calls originating from outside the licensed area can be blocked. Similarly, knowledge of the originating address can be used to selectively limit concurrent users during periods of peak load. There will be situations, however, in which passwords will be necessary. An example is a legitimate user who is trying to access a database from outside a licensed area, such as a faculty member on sabbatical. In this case, the user could bypass that address-checking algorithm by identifying him or herself with a password. Using address checking in general, and IDs only when strictly required, allows the number of users IDs to be reduced, as well as supports any user policies that address prioritized access.

End-user training

As libraries become network resources, the electronic environment users will face every time they login to find information will be very complex. There will be a need to instruct users in the use of network information resources as there has been for traditional paper-bases and microform-based materials. Teaching users about information systems, or bibliographic instruction, is a well-established domain of library work and is considered to be an important component in helping users get the most out of the library (Katz, 1987). Most libraries have programs to instruct users in use of libraries, however, these courses will have to be extended to include topics concerned with information systems that are beyond the bounds of the library. The Mann Library at Cornell University, for example, has a large instructional program that focuses on computer and telecommunications technologies with applications to information retrieval and management (Olsen, 1990). If users are to use the "virtual libraries" that research networks will make possible, then they must be taught the skills needed to use them effectively.

Among the skills that will be needed to deal with the new electronic environment are (Olsen, 1990:231):

- searching bibliographic, full-text, numeric, and directory files
- using information services on CD-ROM as well as mainframe computers
- using telecommunications systems and software
- searching computerized databases on remote systems
- downloading information to floppy disk, and organizing, manipulating, and analyzing this information on a microcomputer using citation-management, spreadsheet, and word-processing software
- using online current awareness services
- accessing library holdings records using the online catalog.

Other skills specific to network use are:

- the ability to navigate in a research network environment including remote login procedures such as TELNET
- the ability to log into specific systems once terminal connection has been completed.
- the ability to use network directories as they develop.

Formal end-user training may be of even greater importance in a networked environment because much of the activity will be done in isolation; that is, the network user will not be in the library where help is usually immediately available from reference staff (Machovec, 1989). These "invisible patrons" will need instruction that will provide them with a basic understanding of searching and of networking concepts. Such basic knowledge can then be effectively supplemented by paper-based instruction sheets and online help.

A secondary issue is that until Z39.50 or SR are widely implemented, libraries may also find themselves faced with the problem of teaching users about *all* types of information systems interfaces that may appear on a research network, such the online catalogues of other institutions. Libraries will have to establish policies with regard to how much of this burden they are willing to take on as the number of information systems available on research networks multiply.

5.2.6 Impact upon staffing

Another sphere upon which the shift to a research network environment will have an impact is library staffing. The impact will be felt primarily on three fronts: 1) a need for a more skilled staff, 2) a need for new positions and designations to deal with the network environment 3) and possible need for increases in the number of staff because of greater activity in areas such as interlibrary loan and end-user training.

Staff skills

The use of research networks will require librarians to have an increased knowledge of, and skills pertaining to, computers, information, and telecommunications technologies. These technologies are conjoined in research networks forming a complex environment that librarians must be able to negotiate to perform library-related and user-related functions. The basis of many of these skills will have already been acquired as earlier innovations, such as integrated library systems, online information retrieval, optical databases, and bibliographic networks, have become commonplace in libraries. These present skills will have to be augmented and adapted to provide those needed for the use of research networks. The types of skills that will be needed can be classified in two areas, general computer skills, and network-oriented skills.

General computer skills These skills include a working knowledge of hardware, software, and programming. Specific skills include:

- the ability to select, set-up, and maintain microcomputers, workstations, and mainframes (Olsen, 1990)
- the ability to mount and maintain tapes of bibliographic and other types of data files
- the ability to reformat and extract sub-sets of data files (Olsen, 1990)
- a familiarity with a variety of software packages such as word processing, database, database construction, spreadsheet, communications, and statistical programs (Olsen, 1990).

Network-oriented skills These skills are those with which library staff may be least familiar. While staff have been performing tasks associated with the skills outlined above for some time, connecting library systems to networks is a relatively new endeavor. Some knowledge and skills that library staff who will be working within this environment should possess include:

- knowledge of network-specific technical equipment such as routers, and PADS, as well as gateway products for turnkey systems
- a working knowledge of networking and network protocols. Librarians may not need to be telecommunications experts, but some knowledge of networking and network protocols will be advantageous

- knowledge of specific applications such as electronic mail, file transfer, and remote login. A good example is the nature and operation of an *anonymous* FTP

- knowledge of LAN technologies including selection, set-up, maintenance, and trouble shooting

- knowledge of address spaces for use in email and in access control.

Developing a staff with these types of skills will require both training, through programs, workshops, conferences, courses, and seminars, as well as by recruiting new staff.

New positions

As library staff begin to develop skills and become responsible for tasks that are widely divergent from those involved in traditional domains, new positions and designations will have to be developed, or old ones re-oriented, to accommodate these changes. Mann Library at Cornell University, for example, has created a number of new positions within their library that reflect this changing environment (Olsen, 1990). Among them are:

- ***Head, Information Technology Section.*** Responsibilities associated with this position are the creation and maintenance of the computing and telecommunications environment within the library. Creation of a position such as this, of course, would depend upon the organizational arrangements made with the campus computing department (see 5.2.4)

- ***Computer Files Librarian.*** This person would be responsible for the selection and service associated with electronic, non-bibliographic data files

- ***Information Literacy Specialist.*** This person's role is to create and teach a curriculum in information literacy. This will become increasingly important as the networking environment becomes more complex

- ***Programmer, Public Services.*** This position includes developing custom software for accessing files made available through the library.

Other conceivable positions related specifically to networks are:

- ***Library LAN administrator.*** A person in this position would be responsible for the selection, set-up, and maintenance of a local area network within the bounds of the library

- ***Network Specialist.*** This person would be responsible for the connection to the network external to the library.

Increased demand for services

Another impact of the use of research networks will be a potential increase in the demand for library services related to information provision. These increases will be the result of characteristics of the new information environment; that is, 1) there will be a larger user group resulting from increased connectivity, and 2) more diverse groups will be targeted through a wider range of information resources.

Although there is little experience as yet with operating a library in a networked environment on a national scale, studies focusing on the effects of tape-mounting bibliographic databases and, in some cases, making them available over campus-wide networks provide some indication of the trends that may be expected. These configurations provide a wider variety of databases directly to remote workstations, and will be similar, though on a smaller scale, to what is proposed for research networks. Examples of systems that provide this type of increased access are the University of California's MELVYL MEDLINE system (Lynch and Berger, 1990), Arizona State University's CARL system (Machovec, 1989), and the M(I)DAS system of Loyola University of Chicago (Bakowski and Moeckel, 1990). Each of these university systems have experienced increased demand for user services as a result of providing greater access to information. The predominant areas affected are reference services and interlibrary loan.

Reference service At these example institutions, reference service departments experienced a dramatic rise in the number of questions asked. The University of California found that reference desk activity "increased substantially" as a result of making MEDLINE available over the campus-wide network (Lynch and Berger, 1990:381). Some UC system libraries were, in fact, forced to limit access to the database to only in-library terminals because of inadequate staff to handle the increased demands. Similar increases were found at Loyola University of Chicago when access was provided to three Wilsonline databases, the *Social Sciences Index*, the *Humanities Index*, and the *Business Periodicals Index*. During a three month period in 1989, the total number of reference questions asked at Loyola increased by 42.6 percent over the same period in the previous year (Bakowski and Moeckel, 1990). If the questions answered by student assistants are included, the total rises to 56.1 percent. Question types included technical questions about the operation of the system, mechanical questions about equipment, and questions about developing complex search strategies. In addition, because users now had greater access to literature found in journals, the increase in reference activity spread to periodical information desks. These desks experienced an increase in questions of a combined total of 82.3 percent. While these data are not based on research network use, per se, they do suggest that increased reference activity will be a general result of wide-spread access to diverse information resources, such as those that will be available through research networks.

Interlibrary Loan ILL rates also rose at the sample universities as a result of increased access to the bibliographic databases. ILL staff in the CARL system, which mounted nine bibliographic databases, experienced a dramatic rise in ILL requests (Machovec, 1989). Requests for photocopies between the campuses of Loyota increased by 150 percent, and requests to libraries outside the Loyota system increased by 21 percent (Bakowski and Moeckel, 1990). Making a wide range of information resources available through research networks will

undoubtedly result in similar increases in ILL activity. As advanced research network services develop, however, and ILL processing becomes more automatic, these demands may be alleviated.

End-user Training Another area that will require increased staff is end-user training (this issue has been discussed at length above, but bears mention here as well). More extensive training courses will require more librarians to develop and teach them.

The increased demand for library services experienced by the above university systems suggests that providing similar services over research networks will also increase demand. These increases may be even more profound as the range of information resources available on research networks, as well as the size of the user group, will surpass those in the example universities discussed above. Libraries must carefully consider how they will accommodate these demands as they connect to research networks.

5.3 Outlook for Developing Nations

Research networks, because they can provide remote access to information, have the potential to deliver to developing nations information resources that are considered essential to their economic and social development. Having recognized the link between science, technology, and economic prosperity, many developing nations have embarked upon programs to organize domestic information systems that collect, process, and disseminate scientific and technological information (Gassol de Horowitz, 1988). These systems are considered to be necessary to improve the scientific and technological base of non-industrialized economies. Even where such systems have been developed, however, they have had little impact upon the growth of science, technology and social development because of numerous constraints to the free flow of information (Dubey, 1986).

The hindrances to information flow are many:

- ***publication systems of developing nations are extremely poor with very low output.*** This low output causes over reliance on advanced countries for primary sources of scientific and technological information

- ***document supply systems are highly inefficient.*** There are few book suppliers in developing countries who accommodate demands for foreign materials

- ***document delivery systems are encumbered by the distance between document suppliers and consumers.*** This distance can often result in delays of four to six months for documents to arrive by surface mail. It also acts to increase the overall cost of the document, as well as the probability that it will be lost in transit

- ***foreign exchange restrictions.*** Foreign exchange restrictions limit the delivery of materials directly from producing countries

- ***the general increase in the costs of publications.*** These increases are especially difficult for developing nations with poor economies

- ***increased costs as a result of photocopying.*** The increasing use of photocopying and other electronic reproduction technologies in industrialized nations, and the concomitant decline in sales, have further increased the price of materials. This has further eroded the purchasing power of developing nations

- ***restrictive copyright laws.*** The increasing use of photocopying and other electronic reproduction technologies has caused copyright law to become more restrictive. Copyright restrictions tend to limit the flow of information outside its country of origin

- ***restrictions on transborder data flow.*** The tendency of industrialized nations to restrict the transborder flow of information considered to be crucial to national security limits the information available to developing nations.

Improvements through the use of research networks

While not all hindrances to the flow of information to developing nations can be solved through the use of research networks, the new information environment created by these networks has the potential to mitigate some of them. Examples of some problems that have the potential to be alleviated are:

- ***improved publication systems.*** One of the major impacts of research networks will be the acceleration of the scholarly communication cycle (Yavarkovsky, 1990); documents are expected to be created, co-authored, edited, peer-reviewed, stored in final form, and accessed online. All scholars may potentially have access to this system regardless of their physical location. This may provide new opportunities for scientists in developing nations to participate in scholarly communication

- ***improved document delivery systems.*** As electronic document delivery systems develop, the efficiency of overall document delivery should improve. This may provide the opportunity for developing countries to have direct access to documents in industrialized nations; problems of large distances, long delays, shipping costs, and losses of materials may, in this way, be eliminated

- ***alleviate problems due to copyright restrictions.*** As discussed in a previous section, new copyright laws must be developed to accommodate the new electronic and research network environment. Success in this area has the potential to: 1) redress the imbalance between information producers and consumers caused by photocopying and other electronic reproduction technologies thereby promoting fairer prices for materials, 2) ensure that "fair use" is extended to the electronic environment allowing free flow of

information. Electronic document delivery systems that have embedded charging mechanisms may ultimately reduce the cost of individual documents. This may potentially increase the purchasing power of developing nations.

Obstacles to research network use

Although research networks can enhance information access, a necessary pre-condition is, of course, the existence of the networks. There are, unfortunately, many obstacles to the development and use of research networks in developing countries. There are few studies directly focused on the implementation of research networks in developing countries. One study (Matta and Boutros, 1989), however, has examined the use of electronic mail in these areas and has identified a number of barriers to its implementation. Because electronic mail is the most basic and ubiquitous of network services, it can be assumed that these barriers will also apply to the general use of research networks and more advanced services, such as those discussed in Chapter 3. The barriers that have been identified fall into four general categories: technological, economic, political, and sociocultural.

Technological barriers In general, the technological resources required to store and transmit data in less developed nations has been described as "inadequate", and the support mechanisms for its widespread use as "nonexistent" (Matta and Boutros, 1989:61,62). Specifically, technological barriers fall into three subcategories:

- ***lack of the necessary technology to receive, store and transmit data.*** This lack of necessary technology includes outdated computer hardware and software, limited use of microcomputers, few instances of LAN connectivity or connections between remote computer sites

- ***lack of reliable land-based communication systems.*** The telephone system is the medium most frequently used for data transfer in developing nations, but in many places telephone systems are scant or non-existent. Furthermore, the transmission quality that they provide for audio and digital data in these countries is poor. (Packet radio, however, has the potential to provide connectivity solutions in developing nations where land-based communications are absent. Packet radio involves the transmission of packets via radio rather than wires (see Brownrigg, *et al*, 1984)).

- ***lack of sufficient technical support.*** This affects the maintenance and reliability of computer and telecommunications systems. There are shortages of systems analysts, programmers, and computer and communications technicians. Shortages of skilled labour also affect the reliability of the electrical power supply.

Economic barriers Severe economic conditions also restrict the development of research network-related services. These conditions are manifested in a poorly trained labour force, high unemployment, lack of capital resources for investment, and a crushing foreign debt. The purchase of the equipment necessary for computer/telecommunications capabilities is, therefore,

beyond the reach of many less developed nations. Furthermore, the relative cost of manual methods is lower than investing in computer and telecommunications equipment.

Political barriers The political barriers to email, and other network services, stem from both internal and external nationalistic and protectionistic forces. For example, industrialized countries view information, as well as information technologies, as a valuable national resource; the export of information and information technology are often restricted for both economic and national security reasons. The same is true of the importation of such materials; the governments of less developed nations frequently view Western technology as threats to national sovereignty, or at least to the sovereignty of those in power.

Sociocultural barriers Sociocultural barriers involve the problems in adapting specific technologies to indigenous patterns. Electronic mail, and certainly research networks, will have a profound impact on all levels of society. Individually, people will have to overcome fear of computers, both because of the lack of computer literacy, as well as fear of displacement. While these factors do exist in industrialized societies, the fact that computer and telecommunications technologies have not been designed for the particular needs of developing countries may make the impact more severe.

That a telecommunications infrastructure supporting research networks must be created before the information necessary to economic prosperity can be made available, and that it will take economic prosperity to enable developing nations to afford a telecommunications infrastructure, is a vicious circle typical of many development issues. Nevertheless, the ability of research networks to provide widespread access to the world's information resources, and the catalyst to change that this information can provide, will doubtless make their implementation in these nations an important issue in the future.

6. FUTURE INITIATIVES

The evolution of research networks, and their interconnection into large national and international internetworks, is progressing throughout the world very rapidly. This swift pace is illustrated by the recent developments of AARNet in Australia, IXI in Europe, NACSIS in Japan, as well as the plans for JANET II in the United Kingdom and the ISTC Network in Canada. Despite these major advances, the development of national and international research networks is still at a very early stage. Signs of their nascency abound; there is lack of uniformity across the networks in underlying protocols, services offered, user policies, and governance. They also have relatively slow transmission speeds, narrow bandwidth, and limited connections to small, or remotely located institutions (Cisler, 1990). Although many research networks currently offer some services that will be a part of future networks, such as access to databases, supercomputers, specialized software, and libraries, they are limited in their scope, ease of use, and functionality. These embryonic networks are not sufficient to create *virtual libraries* or *information superhighways* of the coming decades, which may be built, in part, from the applications discussed in this paper. To create these advanced networks, a first step will be to extend and upgrade the existing network infrastructure well beyond its current speed and capacity. One major initiative in this area, which acts as a indication of what may be in store for the future, is the planned National Research and Education Network or NREN.

6.1 The National Research and Education Network - NREN

The NREN is a major American initiative to develop a high-speed, high-capacity, widely interconnected research and education network that will serve academia, industry, and government. It forms a central component of *The Federal High Performance Computing Program* (HPC Program), a major governmental effort to stimulate the development of high performance computing systems, advanced software technology, networking, and research and education (Bromley, 1989). The magnitude and sheer momentum of the NREN effort is illustrated by the fact that three pieces of legislation were introduced to both the Senate and the House of Representatives that include titles authorizing the development of NREN. The overall HPC program has been formulated in a senate bill sponsored by Senator Albert Gore (D-TN) called the *National High Performance Computer Technology Act of 1990* (S.1067), accompanied by an identical bill in the House (H.R.3131) tabled by Congressman Doug Walgren. A further Senate bill (S.1976) was introduced on behalf of the Department of Energy also containing provisions relating to the NREN. Funding for the NREN component was approximately $400 million over a five year period. While the Gore bill not pass in 1990, a second bill (S.272) pertaining to the initiative will be introduced in 1991.

Generally, the NREN initiative calls on the National Science Foundation (NSF) to upgrade and expand the existing *Internet* infrastructure to establish a three gigabit/second computing network that will connect every university and major laboratory in the U.S by 1996. To put this gigabit figure into perspective, the speed of existing networks range from 56Kbps to 1.56Mbps, and 45Mbps on the soon-to-be-upgraded NSFNET. The speed of the NREN, by contrast, will exceed this by 2000 times. This will allow, for example, 100,000 text pages to be transferred per second (Cisler, 1990). The NREN will be developed in three stages (HPCP, 1990):

- ***Stage 1*** Upgrade the existing *Internet* to 1.5Mbps. This stage is well underway

- ***Stage 2*** Deliver upgraded network services to 200 to 300 research institutions, using a 45Mbps backbone network. (Current figures outstrip this estimate, however, 45Mbps service is presently being deployed, and ANS suggests that speeds of 600 Mbps may be in place by 1993 (Lynch, 1991)). At the end of this stage, it is expected that every university and major laboratory will be connected through mid-level networks

- ***Stage 3*** Deliver 1Gbps to 3Gbps networking service to selected research facilities, and 45Mbps networking to approximately 1000 sites nationwide. The final structure of this network has not been established, pending further research. Deployment of the gigabit network in not expected to begin until the late 1990's.

The governance of the NREN, like current research networks, has not yet been established. The intention is that governmental support will be phased out once the network is up and running, and that it will thereafter become a commercial network, much like the current NSFNET (HPCP, 1990). The success of the NSFNET endeavor, in fact, may influence the eventual governance structure of the NREN (Jackson, 1990).

It is important to point out that the NREN has not been conceptualized as simply a larger conduit for data transfer, but as an enabling technology for increased information access, research, and education, that will profoundly change society. It has been suggested that the impact of the NREN on American society will be as pervasive as the development of the interstate highway system (Gore, 1990). A key element of the NREN initiative is increased information access. Because information access has been the traditional function of libraries, major library associations have become heavily involved in the conceptualization and development of the NREN as an information network. The focus of this activity is the Coalition of Networked Information.

6.2 The Coalition For Networked Information

In recognition of the link between NREN and the library community, the Coalition for Networked Information (CNI) was formed in March, 1990, by three organizations who have as their focus libraries, information technology, and higher education. They are:

- ***The Association of Research Libraries*** (ARL), an organization of major research libraries in the U.S. and Canada

- ***CAUSE***, the association for the management of information technology in higher education

- ***EDUCOM***, a nonprofit consortium of colleges, universities, and other institutions to facilitate the introduction, use, and management of information technology in higher education (Webster, 1990).

The purpose of CNI is to coordinate the activities of the administrators of both libraries and information technologies in higher education institutions. It will provide the focus for discussions concerning the NREN, and facilitate development of the rich information environment enabled by this technology (CNI, 1990a). Their specific missions are:

- to promote the creation of, and access to, information resources in networked environments in order to enrich scholarship and to enhance intellectual productivity

- to seek to realize the information distribution and access potential of existing and proposed high performance computers and networks that support the research and education activities of a wide variety of institutions and organizations

- to undertake activities, on its own and in partnership with others, that formulate, promulgate, evaluate, and promote policies and protocols that enable powerful, flexible, and universal access to networked information resources

- to direct the combined intellectual, technological, professional, and financial resources of its members according to a shared vision of how the nature of information management is changing, and will continue to change, through the end of the 20th century and into the beginning of the 21st.

(CNI, 1990b:1)

CNI will also address policy issues related to the information provision on the NREN, such as intellectual property rights, standards, licensing, servicing arrangements, charging algorithms and cost recovery fees, and economic models (CNI, 1990a). These missions will be carried out by a task force composed of the institutions that comprise ARL, CAUSE, and EDUCOM. At present, approximately 200 organizations and institutions have committed to join the task force.

7. SUMMARY

Research networks, because they have the ability to support many forms of communication between computers and their users, have the potential to evolve into an information infrastructure that will provide greater capabilities for the management and flow of information. This developing infrastructure will enable a larger and more diverse set of information resources to be accessed by a broader community of users, and promote increased information exchange among them. Contemporary libraries, which increasingly utilize electronic forms of information, as well as computer and telecommunications technology, have an opportunity to both contribute to and benefit from this effort. There are also, however, a number of obstacles to be overcome before end-user applications of research networks can be developed. These benefits and obstacles are summarized below.

Benefits of library use of research networks

The successful implementation of research network applications described in this paper have the potential to benefit both libraries and their users. Applications in the areas of bibliographic information provision and document supply, specifically, may provide the following benefits:

- allow more people to access a wider variety of information resources through distributed information retrieval
- enable information to be accessed in the workplace, regardless of the location of the workplace
- facilitate the use of information systems by eliminating the need to have detailed knowledge of various system interfaces, search conventions, or database formats
- streamline current awareness programs by allowing SDI profiles to be searched automatically across numerous databases, and by transmitting the results through electronic mail
- simplify and enhance the interlibrary loan process by providing a common channel for the exchange of standardized interlibrary loan messages
- accelerate the delivery of documents by providing a mechanism for their delivery directly to end-users in electronic form
- increase the potential for the international exchange of scholarly information.

There are also a number of benefits of technical service applications of research networks. Such applications have the potential to:

- facilitate the direct machine-to-machine exchange of cataloguing records and other data, thereby increasing productivity and reducing costs
- facilitate the creation and maintenance of union catalogues
- support the exchange of business messages involved in the acquisitions process
- increase the potential for the international exchange of cataloguing data.

In addition, there are also overall benefits in the use of research networks. For example, they have the potential to:

- distribute the costs involved in maintaining numerous databases among a number of libraries interconnected through a research network
- reduce costs over the long term and allow libraries to make better use of existing information technology
- increase the profile of libraries as an essential component of research initiatives
- assist libraries in responding to the general trend in publishing toward the automation of all stages of the publishing process.

Obstacles to library use of research networks

A number of obstacles must be overcome, however, before library applications of research networks can be fully implemented. These include not only technical obstacles, but also legal and political obstacles that will arise during the creation of information networks. These obstacles include:

- the inability of the current generation of library systems to connect to research networks
- the inability to access research networks from *within* library applications
- the lack of standards for data structures such as those pertaining to patron information as well as to abstracting and indexing databases
- the lack of widespread use of standard document architectures, such as ODA
- the lack of OSI products, or OSI products that offer the full functionality described in the standards

- the absence of an online directory to help users find information and navigate through the networks
- the lack of an overall governance structure that will represent all relevant points of view, such as those of individual research networks and existing library networks
- the failure of copyright laws to keep pace with innovations in information technology, and a concomitant lack of legal mechanisms to compensate information producers fairly, while avoiding undue restrictions upon information consumers
- the incompatibility between restrictions on commercial use and the need to deal with commercial concerns to develop advanced network services
- the need to renegotiate licensing arrangements for restricted databases allowing them to be used in a network environment
- the need to establish some form of communication and cooperation between libraries and academic computing centres
- the need to modify existing user policies and to create new policies to control access to limited system resources as well as to databases restricted by licensing agreements
- the need to develop wider staff skills to deal with a more complex networking environment
- the need to cope with increased demands for services resulting from enhanced information access.

Role of the library

Because libraries have had long experience in the organization, management, and dissemination of information, they have a unique opportunity help determine the future shape of information networks. Resting on this long tradition, however, is not enough to make libraries full partners in this endeavor. Librarians must actively (Webster, 1990):

- participate in forums on the evolving information infrastructure
- evaluate new technologies and explore cooperative projects
- assist in the development of directories of information resources, including the development of gateways to these resources

- assist in the development and promotion of standards, and in the migration of existing systems to OSI-based systems

- educate users in the use of the developing information infrastructure.

The complex information network that is currently evolving has the potential to transform many sectors of society that utilize information. Libraries, as mediators of information access, will be among those institutions that will be most profoundly altered as the use of this infrastructure becomes as commonplace and as effortless as the telephone system. The implementation of the applications, protocols, and standards described in this paper have the potential to form, in part, the basis for this new infrastructure.

References

AARNet. (1990). *Australian Academic and Research Network Resource Guide.* 9th October 1990.

ACOSITG. (1987). Academic Community OSI Transition Group. *Transition to OSI Standards (White Book).*

Andre, P.Q.J., Eaton, N.L. and Zidar, J.A. (1989). Scanning and digitizing technology employed in the National Agricultural Text Digitizing Project. In *Proceedings of the Conference on Application of Scanning Methodologies in Libraries.* D.L Blamberg, C.L. Dowling, and C.V. Weston (eds.). Beltsville, Maryland: National Agricultural Library. pp. 61-73.

Arms, C.R. (1988a). Protocols and standards. In *Campus networking strategies.* C. Arms (ed.). Bedford, Mass.: Digital Press. pp. 255-269.

Arms, C.R. (1989a). Libraries and electronic information: the technological context, Part One. *EDUCOM Review, 24*(2), 38-43.

Arms, C.R. (1989b). Libraries and electronic information: the technological context, Part Two. *EDUCOM Review, 24*(3), 34-43.

Arms, C.R. (1990a). *Campus Strategies for libraries and electronic information.* EDUCOM Strategies Series on Information Technology. Digital Press 1990.

Arms, C.R. (1990b). A new information infrastructure. *Online, 14*(5), 15-22.

Arms, C.R. (1990c). Using the national networks: BITNET and the *Internet. Online, 14*(5), 24-29.

Arvillas, A. (1990). The Greek ARIADnet Network. *IES News, 31*, 18.

Avram, H. (1988a). Building a unified information network. *EDUCOM Bulletin*, 23(4), 11-14.

Avram, H. (1988b). LSP and library network services in the future. *EDUCOM Bulletin, 23*(2/3), 52-58.

Avram, H. (1989). Copyright in the electronic environment. *EDUCOM Review, 24*(3) 31-33.

Bakowski, V.B., and Moeckel, L.E. (1990). Mounting local tapes: the NOTIS interface combines turnkey ease and adaptability. *Online, 14*(4), 38-42.

Barron, B. (1990). *UNT's On-line Library Handout.* University of North Texas.

BIBSYS. (1990). *BIBSYS System Description.* Dragvoll, Norway: BIBSYS.

Boss, R.W., and McQueen, J. (1983). *Document delivery in the United States.* Information Systems Consultants, Inc.

Bromley. D.A. (1989). *The Federal High Performance Computing Program.* Executive Office of the President. Office of Science and Technology Policy. U.S.A.

Brownrigg, E.B., Lynch, C.A., and Pepper, R.L. (1984). Packet radio for library automation. *Information Technology for Libraries, 3*(3), 229-244.

Buckland, M.K. and Lynch, C.A. (1988). National and international implications of the Linked Systems protocol for online bibliographic systems. *Cataloguing and Classification Quarterly, 8*(3/4), 15-33.

Buitelaar, W.G. (1990). SURFnet and PICA in open library networking. *IES News*, 29, 4-5.

Buxton, Andrew. (1988). JANET and the librarian. *The Electronic Library, 6*(4), 250-263.

Cahan, M.A. (1989). Optical character readers/text scanners: a market analysis. *Information Technology and Libraries*, 8(2), 186-196.

Callon, R., Hagens, R. and Nitzan, R., (1990). *Internet OSI integration, coexistence and interoperability issues.* Draft, *Internet* NIC file.

Cartwright, D., Gilbert, S.W., and Lyman, P. (1989). Beyond intellectual property. *EDUCOM Review, 24*(2), 7-8.

Cawkell, A.E. (1987). A survey of image processing and document delivery technology. *British Library Research Paper, 28.* British Library Board.

CDNnet Reports, 5, 1989, 4.

Chamberlain, C.E. (1985). Automating acquisitions: a perspective from the inside. *Library Hi Tech, 3*(3), 57-66.

Chapman, E. (1990). (RLIN Information Center). Personal communication.

Chapman, E. (1991). (RLIN Information Center). Personal communication.

Cisler, S. (1990). NREN: the National Research and Education Network. *LITA Newsletter, 11*(2), 1-2.

Cline, N.M. (1990). Information resources. *Educom Review*, *25*(2), 30-34.

CNI. (1990a). *ARL, CAUSE, and EDUCOM announce the Coalition for Networked Information.* News release, March 16, 1990. Coalition for Networked Information.

CNI. (1990b). Program strategy. *Report on the June 1990 Organizational Meeting.* Coalition for Networked Information.

Comer, Douglas E. (1988). *Internetworking with TCP/IP: Principles, Protocols, and Architecture.* Englewood Cliffs, N.J.

Cooper, R. (1990). *The JANET Network.* The Computer Board for Universities and Research Councils.

Cornish, G.P. (1990). *The conflict between copyright and document supply: real or imagined?* Paper presented at the 56th IFLA General Conference in Stockholm, Sweden, August, 1990.

Costers, L. (1990). The new PICA library network. *Division of Management and Technology.* Session paper presented at the 56th IFLA General Conference, Stockholm, Sweden, August 18-24, pp. 13-16.

Dawson, F. and Nielsen, F. (1990). ODA and document interchange. *UNIX Review*, *8*(3), 50-56.

DeCandido, G.A. and Rogers, M. (1990). "Virtual library" promulgated by library/education coalition. *Library Journal*, April 15, 14.

Dekkers, Makx. (1990). Pica and networking between library systems. In *Bibliographic Access in Europe: 1st International Conference.* Lorcan Dempsey (ed.). Aldershot, U.K.: Gower Publishing Company. pp. 136-138.

Dempsey, L. (1990). Bibliographic access: patterns and developments. In *Bibliographic Access in Europe: 1st International Conference.* Lorcan Dempsey (ed.). Aldershot, U.K.: Gower Publishing Company. pp. 1-29.

Dempsey, L. (1991a). (Centre for Bibliographic Management, University of Bath) Personal communication.

Dempsey, L. (1991b). *Libraries, networks and OSI: a review with a report on North American developments.* Bath, U.K.: UK Office for Library Networking.

Denenberg, R. (1988). Standard Network Interconnection Protocols. In *The Linked Systems Project: networking tool for libraries*. J.G. Fenly and B. Wiggins (eds.). Dublin, Ohio: OCLC Online Computer Library Center.

DiDio, Laura. (1989). Users rely on TCP/IP, wait for interoperable OSI wares. *Network World*, Nov. 6, 25,28.

Dubey, Y.P. (1986). Information poverty: a Third World perspective. In *Unequal Access to Information Resources: Problems and Needs of the World's Information Poor*. J.P. Lang (ed.). Ann Arbor, MI: Pierian Press. pp. 47-54.

Fisher, Sharon. (1989). OSI takes on TCP/IP. *Unix World*, Feb. 74,75,79,80.

Fjallbrant, N. (1988). Document delivery by satellite. *International Forum on Information and Documentation*, *13*(3), 21-25.

Fuchs, Ira. (1988). Research networks and acceptable use. *EDUCOM Bulletin*, *23*(2/3), 43-48.

Gassol de Horowitz, R. (1988). *Librarianship: a Third World Perspective*. New York: Greenwood Press.

Goldstein, M.L. (1986). Meet the new entrepreneur: lawyers must sort out who owns and is liable for electronic data. *Industry Week*, April 28, 65.

Gore, A. (1990). Superhighways for computing. *Computer World*, September 3, 21.

Gunderson, R. (1991). (BYBSYS Project Manager). Personal communication.

Heaney, Henry. (1989). Networking and interlending in the United Kingdom. *Interlending and Document Supply*, *17*(2), 16-18.

Helfer, D. (1990). EPIC: the review. *Database Searcher*, Nov., 15-23.

Hickling Science and Technology Division. (1990). *Feasibility study of a national high speed communications network for research, development, and education*. A report submitted to Industry, Science and Technology Canada.

Huston, Geoff. (1990a). The Australian Academic and Research Network. *Australian Association of Research Libraries*, May, 1-11.

Huston, Geoff. (1990b). (AARNet Network Technical Manager). Personal communication.

IES News. (1990a). IXI: a European X.25 network pilot project. *IES News*, *26*, 7-8.

IES News. (1990b). The German network for science. *IES News, 27*, 1-2.

IES News. (1990c). Networking in the former German Democratic Republic. *IES News, 30*, 26-27.

Information World Review. (1991). Academic network starts fixed-price searching. *Information World Review*, January (55), 3.

Ishida, Haruhisa. Research networks in Japan.(1988). In *Computer Technologies for the 80's*. J. Raviv (Ed.). North Holland: Elsevier Science Publishers, 416-418.

Jackson, Kelly. (1989). Commercialization of the *Internet*. *Communications Week*, Oct 2, 25-28.

Jackson, Kelly. (1990). Opening the *Internet*. *Communications Week*, 281, 1+.

Jacob, M.E. (1990b). Libraries and national library networks. *Bulletin of the American Society for Information Science*, 16(5), 8-9.

Johnson, E.R. (1987) Introduction: Management issues in the networking environment. *Journal of Library Administration*, 8(3-4), 5-6.

Katz, W. (1987). *Introduction to reference work*. New York: McGraw-Hill.

Kaufman, P.T. (1990). Columbia University. In *Campus Strategies for Libraries and Electronic Information*. Caroline Arms (ed.). U.S.A.: Digital Press Corporation. pp. 194-217.

Kibbey, Mark, and Evans, Nancy H. (1989). The network is the library. *EDUCOM Review, 24*(3), 15-20.

Kinney, T. (1988). *Toward telecommunications strategies in academic and research libraries: ten case studies of decision-making and implementation*. Washington, D.C.: Office of Management Services, A.R.L.

Landweber, L.H., and Jennings, D.M. (1986). Research computer networks and their interconnection. *IEEE Communications Magazine. 24*(6), 5-17.

Larsen, R.L. (1990). The colibratory: the network as testbed for a distributed electronic library. *Academic Computing*, Feb., 22-23, 35-37.

Lauck, Anthony G., Oran, David R., and Perlman, Radia J. (1986). Digital Network Architecture Overview, *Digital Technical Journal, 3*, 10-24.

Lowry, C.B. (1987). Ownership of bibliographic data and its importance to consortia. *Journal of Library Administration*, *8*(3/4), 69-84.

Lynch, C.A. (1988). The evolution of "real" networks. In Networking in transition: current and future trends: a forum edited by Chris Sugnet. *Library Hi Tech*, *6*(4), 101-119.

Lynch, C.A. (1989a). Library automation and the national research network. *EDUCOM Review*, *24*(3), 21-26.

Lynch, C.A. (1989b). Linking library automation systems in the *Internet*: functional requirements, planning and policy issues. *Library Hi Tech*, *7*(4), 7-18.

Lynch, C.A. (1990). The growth of computer networks: a status report. *Bulletin of the American Society for Information Science*, *16*(5), 10-11.

Lynch, C.A. (1991). (Director of Library Automation, University of California). Personal communication.

Lynch, C.A. and Berger, Micheal G. (1990). The UC MELVYL MEDLINE system: a pilot project for access to journal literature through an online catalog. *Information Technology and Libraries*, *8*(4), 371-383

Lynch, C.A. and Preston, C.M. (1990). *Internet* access to information resources. In *Annual Review of Information Science and Technology*, *25*. Martha E. Williams, (ed.). Amsterdam: Elsevier Science Publishers. pp. 1-47.

Lyons, P. (1990). *Copyright considerations: imaging and document conversion.* Preconference paper presented at the Library and Information Technology Association, Chicago, Illinois, June 22.

MacColl, J.A. (1990). Network services to libraries. *IES News*, 27, 8.

Machovec, G.S. (1989). Locally loaded databases in Arizona State University's online catalog using the CARL system. *Information Technologies and Libraries*, *8*(2), 161-171.

MacKinnon, D., McCrum, W., and Sheppard, D. (1990) *An Introduction to Open Systems Interconnection.* New York: Computer Science Press.

Mandelbaum, R. (1990). Free-market networking. *Educom Review*, *25*(1), 10-13, 66.

Marshall, P. (1989). *Onet: the first year.* Paper presented at NetNorth/Canadanet Conference, Montreal.

Martin, James. (1988) . *Data Communication Technology.* Englewood Cliffs, N.J.

Martin, S.K. (1986). *Library networks: libraries in partnership.* White Plains, N.Y.: Knowledge Industry Publications.

Mathews, W.D. (1979). Impact of technology on governance. In *Structure and governance of library networks.* Allen Kent and Thomas Galvin (eds.). New York: Dekker, pp. 121-140.

Matta, K.F. and Boutros, N.E. (1989). Barriers to electronic mail systems in developing countries. *The Information Society, 6,* 59-68.

McCallum, S.H. (1988). Overview of the Linked Systems Project. In *The Linked Systems Project: networking tool for libraries.* J.G. Fenly and B. Wiggins (eds.). Dublin, Ohio: OCLC Online Computer Library Center.

McGill, M. and Racine, D. (1990). OCLC: Online Computer Library Center. In *Campus Strategies for Libraries and Electronic Information.* Caroline Arms (ed.). U.S.A.: Digital Press Corporation. pp. 37-56.

Meijer, Anton. (1988). SNA Trends. *Networks for the 1990s,* Ray Reardon (ed.). New York: Wiley.

Merilees, Bobbie (1989) Canadian library integrated systems: third annual survey. *Canadian Library Journal,* June 1989, 181-187.

MIC News. (1990). SUNET login procedure. *MIC News, 3,* 7.

Miller, N.P. and Blumenthal, C.S. (1986). Intellectual property issues. In *Toward a Law of Global Communications Networks.* Anne W. Branscomb (Ed.). New York: Longman. pp. 227-237.

Molholt, P. (1985). On converging paths: the computing center and the library. *Journal of Academic Librarianship, 11*(5), 284-288.

Naito, H. (1989). Japan's national Science Information System. *Outlook on Research Libraries,* 11(11), 1-3.

Neff, Raymond K. (1985). Merging libraries and computers centers: manifest destiny or manifestly deranged? *EDUCOM Bulletin,* (Winter), 8-16.

New Scientist, (1990). Soviet connection. *New Scientist, 128*(1737), 23.

NSF Network Service Center. (1990). *Internet Resource Guide.* BBN Systems and Technologies Corporation. Cambridge, MA. (Electronic document).

NSFNET Flyer. (1989). *NSFNET - The National Science Foundation Computer Network for Research and Education.* Information pamphlet.

Olausson, C. (1990). Automation of academic libraries and data communications development in Sweden. In *Bibliographic Access in Europe: 1st International Conference.* Lorcan Dempsey (ed.). Aldershot, U.K.: Gower Publishing Company. pp. 128-135.

Olausson, C. (1991). (Librarian and Coordinator of Academic Library Development, Stockholm, Sweden). Personal communication.

Olsen, J.K. (1990). Cornell University, Mann Library. In *Campus Strategies for Libraries and Electronic Information.* Caroline Arms (ed.). U.S.A.: Digital Press Corporation. pp. 218-242.

Plassard, M-F, (1989). The impact of new technology on document availability and access. *Interlending and Document Supply, 17*(1), 3-10.

Plassard, M-F., and Line, M. (1988). *The Impact of New Technology on Document Availability and Access.* (expanded version). Wetherby, West Yorkshire: British Library Document Supply Centre.

Popper, P. (1990). The German network for science. *IES News*, 27, 1-2.

Post, W. and Sessions, J.A. (1986). Academic institutions and information services: the position of the library. *Library Hi Tech News, 29*, 7-9.

Potter, W.G. (1989). Expanding the online catalog. *Information Technologies and Libraries, 8*(2), 99-104.

Quarterman, John S. (1990). *The Matrix: Computer Networks and Conferencing Systems Worldwide.* Bedford, Mass.: Digital Press.

Radlett, A.L. (1988). Office document architecture. In *Networking Technology and Architectures.* Pinner, Middx, UK: Online Publications, 299-309.

Richards, D. (1990). The Research Libraries Group. In *Campus Strategies for Libraries and Electronic Information.* Caroline Arms (ed.). U.S.A.: Digital Press Corporation. pp. 58-94.

Richards, D., and Lerche, C. (1989). Evolution of the Research Libraries Information Network, *Library Hi Tech, 25*(7), 29-36.

Robinson, P.J. and Strasen, S.M. (1989). Standard Page Description Language. *Computer Communications*, *12*(2), 85-92.

Rose, M.T. (1989). *The Open Book: a Practical Perspective on OSI*. Englewood Cliffs, NJ: Prentice Hall.

Rose, M.T. and Cass, D.E. (1987). OSI transport services on top of the TCP/IP. *Computer Networks and ISDN Systems*, *12*, 159-173.

Rose, M.T. and Schoffstall, M.L. (1990). *An introduction to a NYSERNet white pages pilot project*. NYSERNet, Inc.

Schein, A. (1989). Optical storage and OCR -- key components of automated information management systems. *Optical Information Systems*, *9*(1), 9-15.

Shaughnessy, T.W. (1987) Management perspectives on network membership. *Journal of Library Administration*, 8(3-4), 7-14.

Siciunas, E. (1990). *CAnet*. Paper presented at Networking '90: The Canadian Networking Conference, Victoria, June 13.

Soini, Annti. (1990). LINNEA - Library information network for Finnish academic libraries. In *Bibliographic Access in Europe: 1st International Conference*. Lorcan Dempsey (ed.). Aldershot, U.K.: Gower Publishing Company. pp. 113-120.

Spratt, E. Brian (1989). Networking Developments in the UK Academic Community, *Proceedings of the International Conference on Information Network and Data Communication* (Ronneby Brunn, Sweden: 11-14 May 1986). Amsterdam: Elsevier.

St. George, A. and Larsen (1990). *Internet-accessible library catalogues and databases*. November, 1990. (Document available over BITNET: send the message GET INTERNET LIBRARIES to LISTSERV@UNMVM.BITNET).

Stone, P. (1990a). *The development of library and information services on academic and research networks*. European Conference on Library Automation and Networking, May 1990, 5-15.

Stone, P. (1990b). Developing networked services for libraries: the U.K. experience. *Computer Networks and ISDN Systems*, *19*, 343-349.

Stone, P. (1990c). *JANET: A Report on Its Use for Libraries*. British Library Research Paper #77. The British Library Research and Development Department. The British Library Board.

Stone, P. (1990d). (University of Sussex). Personal communication.

Stone, P. (1991). (University of Sussex). Personal communication.

Swain, S.L. (1990). Message Handling Systems and Electronic Data Interchange: an introduction to converging standards for electronic messaging. *IFLA Journal*, *16*(2), 204-214.

Swain, S.L., and Tallim, P., (1990). X.400: the standard for Message Handling Systems. *Library Hi Tech*, *8*(4), 43-56.

Tuck, B. (1989). *OSI and Library Services: Report to the British Library of a 3 Month Study.* London: British Library.

Turner, F. (1990b). The Interlibrary Loan Protocol: an OSI solution to ILL messaging. *Library Hi Tech*, 8(4), 73-82.

UDT Newsletter. (1990) Electronic Data Interchange (EDI) standards. *UDT Newsletter*, *13*, 3.

Walker, F.L. (1989). Issues in document conversion. In *Proceedings of the Conference on Application of Scanning Methodologies in Libraries.* D.L Blamberg, C.L. Dowling, and C.V. Weston (Eds.). Beltsville, Maryland: National Agricultural Library. pp. 45-59.

Wallberg, H. (1990). *SUNET: the Swedish University Network.* Paper distributed at NORDUNET '90.

Walters, E.M. (1987). The issues and needs of a local library consortium. *Journal of Library Administration*, 8(3-4), 15-29.

Walton, R.A. and Bridge, F.R. (1988). Automated systems marketplace 1987: maturity and competition. *Library Journal*, April, 33-44.

Webster, D. (1990). *Announcement of Coalition.* Information packet distributed to ARL directors March, 20, 1990.

Wesley-Tanaskovic, I. (1990). Computerized academic communications networks in Africa. *IES News*, 27, 8.

Wheeler E.F., and Ganek, A.G. (1988). Introduction to Systems Application Architecture, *IBM Systems Journal*, *27*(3), 250-263.

Woodsworth, A. (1988). Libraries and the Chief Information Officer: implications and trends. *Library Hi Tech*, *6*(1), 37-44.

Wu, Gilbert. (1989). SGML theory and practice. *British Library Research Paper 68.* British Library Research and Development Department.

Yavarkovsky, J. (1990). *Vision statement: a university-based electronic publishing network.* Association of Research Libraries ARL/CAUSE/EDUCOM Coalition on Networked Information Task Force Meeting.

Relevant Standards

The list is not complete, but includes the ISO (International Organization for Standardization) standards mentioned in the study as well as some other relevant standards. In addition, it does not include most standards at the Physical layer, nor does it include many industry-specific Application layer standards other than for library-oriented applications.

No distinction has been made in this list between the various stages of standards development; thus both standards and addenda could be at the draft proposal or draft international standard or full international standard stage.

CCITT recommendation numbers have been included where they are available

OSI Reference Model

7498-1	Basic Reference Model (X.200)
7498-2	Security Architecture
7498-3	Naming and Addressing
7498-4	OSI Management Framework
7498 AD1	Connectionless Operation
7498 AD2	Multipeer Data Transmission
TR 8509	Service Conventions (X.210)

Application Layer

9545	Application Layer Structure
AD1	Connectionless Operation

Association Control Service Element(ACSE)

8649	Service Definition (X.217)
8650	Protocol Specification (X.227)
10035	A-UNIT-DATA Protocol

Commitment, Concurrency and Recovery Service Element (CCRSE)

9804	Service Definition (X.237)

9805	Protocol Specification (X.247)

Reliable Transfer Service Element (RTSE)

906-1	Model and Service Definition (X.218)
9066-2	Protocol Specification (X.228)

Remote Operations Service Element (ROSE)

9072-1	Model, Notation and Service Definition (X.219)
9072-2	Protocol Specification (X.229)
10148	Basic Remote Procedure Call Using OSI Remote Operations

Common Management Information Service Element (CMISE)

9595	Service Definition
9596	Protocol Specification

File Transfer, Access and Management (FTAM)

8571-1	General Introduction
8571-2	Virtual Filestore Definition
8571-3	File Service Definition
8571-4	File Protocol Specification
8571-5	PICS Proforma

Virtual Terminal (VT)

9040	Service Definition - Basic Class
9041	Protocol Specification - Basic Class

Job Transfer and Manipulation (JTM)

8831	Concepts and Services
8832	Specification of the Basic Class Protocol

Directory

9594-1	Overview of Concepts, Models, and Services (X.500)
9594-2	Information Framework (X.501)
9594-3	Access and System Services Definition (X.511)
9594-4	Procedures for Distributed Operation (X.518)
9594-5	Access and System Protocols Specifications (X.519)
9594-6	Selected Attribute Types (X.520)
9594-7	Selected Object Classes (X.521)
9495-8	Authentication Framework (X.509)
9495-9	Directory Information Tree Structure and Naming
9495-10	Replication and Knowledge Management

Message Oriented Text Interchange System (MOTIS)

10021-1	System and Service Overview (X.400)

10021-2 Overall Architecture (X.402)
10021-3 Abstract Service Definition Conventions (X.407)
10021-4 Abstract Service Definition and Procedures (X.411)
10021-5 Message Store - Abstract Service Definition (X.413)
10021-6 Protocol Specification (X.419)
10021-7 Interpersonal Messaging System (X.420)
Conformance Testing (X.403)
Encoded Information Type Conversion (X.408)

Remote Database Access (RDA)
9579 Generic Remote Database Access Service and Protocol

Transaction Processing (TP)
10026-1 Model
10026-2 Service Definition
10026-3 Protocol Specification

Interlibrary Loan (ILL)
10160 Service Definition
10161 Protocol Specification

Search and Retrieve (SR)
10162 Service Definition
10163 Protocol Specification

Presentation Layer Standards

8822 Connection Oriented Presentation Service Definition (X.216)
8823 Connection Oriented Presentation Protocol Specification (X.226)
9576 Connectionless Presentation Protocol Specification
8824 Abstract Syntax Notation 1 (ASN.1) Specification (X.208)
8825 Basic Encoding Rules for Abstract Syntax Notation (X.209)

Session Layer Standards

8326 Basic Connection Oriented Session Service Definition (X.215)
8327 Basic Connection Oriented Session Protocol Specification (X.225)
9548 Connectionless Session Protocol

Transport Layer Standards

8072 Transport Service Definition (X.214)
8073 Connection Oriented Transport Protocol Specification (X.224)
8602 Protocol for Providing the Connectionless-mode Transport Service

Network Layer Standards

8348	Network Service Definition (X.213)
8348	AD1 Connectionless-mode Transmission
8348	AD2 Network Layer Addressing
8348	AD3 Additional Features of the Network Service
8473	Protocol for Providing the Connectionless-mode Network Service
9542	End System to Intermediate System Routing Exchange Protocol for Use in Conjunction with the Protocol for the Provision of the Connectionless-mode Network Service
8648	Internal Organization of the Network Layer
8880	Protocol Combinations to Provide and Support the OSI Network Service, Part 1: General Principles (Parts 1-4)
9574	Provision of the OSI Connection-mode Network Service by Packet-mode Terminal Equipment Connected to an ISDN
TR 9575	OSI Routing Framework
TR 9577	OSI Protocol Identification in the Network Layer (X.244)
10028	Definition of the Relaying Functions of a Network Layer Intermediate System

X.25 Packet Level Protocol

8208	X.25 Packet Level Protocol for Data Terminal Equipment (X.25)
8881	Use of the X.25 Packet Level Protocol in Local Area Networks
8878	Use of X.25 to Provide the Connection-oriented Network Service (X.223)
10030	End System to Intermediate System Routing Information Exchange Protocol for Use in Conjunction with ISO 8878
TR 10029	Operation of an X.25 Interworking Unit

Data Link Layer Standards

8886 Data Link Service Definition (X.212)

High Level Data Link Control (HDLC)

3309	High level Data Link Control (HDLC) - Frame Structure
4335	HDLC - Consolidation of Elements of Procedures
7478	Multi-link Procedures (X.25)
7776	HDLC - Description of the X.25 LAPB compatible DTE Data Link Procedures (X.25)
7809	HDLC - Consolidation of Classes of Procedures
8885	HDLC - General Purpose XID Frame Information Field Content and Format

Local Area Networks

10038	Media Access Control (MAC) Bridges
10039	Media Access Control (MAC) Service Definition

8802-1 Local Area Networks, Part 1: Introduction
8802-2 Local Area Networks, Part 2: Logical Link Control
8802-3 Local Area Networks, Part 3: Carrier Sense Multiple Access with Collision Detection - Access Method and Physical Layer Specifications
8802-4 Local Area Networks, Part 4: Token-passing Bus Access Method and Physical Layer Specification
8802-5 Local Area Networks, Part 5: Token Ring Access Method and Physical Layer Specification
8802-7 Local Area Networks, Part 7: Slotted Ring Access Method and Physical Layer Specification

Physical Layer Standards

10022 OSI Physical Layer Service Definition
(numerous connector standards exist, but have not been listed)

Conformance Testing

9646 OSI Conformance Testing Methodology and Framework, Part 1: General Concepts (X.290) (Parts 1-6)

Registration Authorities

9834 Procedures for Specific OSI Registration Authorities, Part 1: General Procedures (Parts 1-6)

Formal Description Techniques

8807 LOTOS - A Formal Description Technique Based on the Temporal Ordering of Observational Behavior
9074 ESTELLE - A Formal Description Technique Based on an Extended State Transition Model

International Standardized Profiles

TR 10000-1 International Standardized Profiles, Part 1: Taxonomy Framework
TR 10000-2 International Standardized Profiles, Part 2: Taxonomy Update, ISP Approval, and Maintenance Procedures
TR 10000-3 International Standardized Profiles, Part 3: Taxonomy of Profiles
TR 10000-4 International Standardized Profiles, Part 4: Directory of Profiles and ISPs

Text and Office Systems

8613-1	Office Document Architecture (ODA) and Interchange Format, Part 1: Introduction and General Principles
8613-2	Office Document Architecture (ODA) and Interchange Format, Part 2: Document Structures
8613-4	Office Document Architecture (ODA) and Interchange Format, Part 4: Document Profile
8613-5	Office Document Architecture (ODA) and Interchange Format, Part 5: Document Interchange Format
8613-6	Office Document Architecture (ODA) and Interchange Format, Part 6: Character Content Architectures
8613-7	Office Document Architecture (ODA) and Interchange Format, Part 7: Raster Graphics Content Architectures
8613-8	Office Document Architecture (ODA) and Interchange Format, Part 8: Geometric Graphics Content Architectures
10033	Recording of Documents Conforming to ISO 8613 on Flexible Disk Cartridges Conforming to ISO 9293
8879	Standard Generalized Markup Language (SGML)
9069	SGML Support Facilities - SGML Document Interchange Format (SDIF)
9070	SGML Support Facilities - Registration Procedures for Public Text Owners
TR 9573	SGML Support Facilities - Techniques for Using SGML
10031-1	Distributed Office Applications, Part 1: Model
10031-2	Distributed Office Applications, Part 2: Referenced Data Transfer

Document Transfer, Access, and Manipulation (DTAM) - Composite Graphics Content Architectures (T.419)

Document Transfer, Access, and Manipulation (DTAM) - Introduction and General Principles (T.431)

Document Transfer, Access, and Manipulation (DTAM) - Service Definition (T.432)

Document Transfer, Access, and Manipulation (DTAM) - Protocol Specification (T.433)

Document Transfer, Access, and Manipulation (DTAM) - Operational Structure (T.441)

APPENDIX A: GLOBAL RESEARCH NETWORK TOPOLOGY

The number of computer networks that span the globe has grown rapidly since the appearance of the first networks in the U.S. in the late 1960's. Although almost impossible to enumerate accurately, the total number of computer networks, of all kinds, currently exceeds hundreds worldwide. That total fluctuates continually as new networks are created, old ones disappear, two existing networks merge into one, or are subsumed under larger metanetworks. Two clear trends emerge, however, that networks will increase in size, and that the varied protocols underlying them will become progressively homogeneous. These trends will lead eventually to the development of a single, completely interconnected world-wide network, much like the telephone network that presently covers the globe (Quarterman, 1990).

The current state of computer networks, however, is still far from this ideal. Rather than a homogeneous system, the reality is a patchwork of networks of differing characteristics. Computer networks can vary in the types of data networks that support them (giving rise to differences in transmission media and transmission speed), in the purpose for which they were created, in the underlying protocols they use, and in the services they offer. They also differ in constituent membership, in how they are administered and funded, in their geographic scope, and in how they are interconnected with other networks. In order to provide a concise, overall picture of this global topology, this appendix contains descriptions of each major network broken down by a number of categories relating to the central characteristics of network systems. These categories are name, purpose, geographic location, technical configuration, governance, financing, membership, and interconnections to other networks. In cases where the information is unavailable, the category is omitted. The result is a collection of summaries briefly outlining each network. The scope of each category is outlined below. For further details about networks and networking in general, please see, *The Matrix*, John Quarterman's (1990) comprehensive book on computer networks. Much information in this chapter is derived from this work, and has been updated and expanded from other sources where possible.

Due to the extremely rapid pace of change characteristic of computer networking, any description of research networks is out of date almost immediately after it is written. Therefore, this section is intended to make librarians aware of the great extent to which research networks are currently used worldwide, and to provide them with a little background information, rather than to provide a complete and up-to-date network directory.

Categories

Name In order to deal with the plethora of acronyms that make up most network names, the meaning of each is stated fully.

Date This is the year the network became operational.

Purpose Networks can be classified according to their central purpose. Research networks, for example, is only one broad category within many network types. Others include military networks, experimental networks, networks set up to share supercomputer resources, networks supporting special interest groups, commercial networks, Public Data Networks (PDNs), networks internal to large companies, and government networks.

The networks selected for discussion in this chapter are those that are in some way related to research, education, or academia. That is, they link together the hosts that reside at government research centres, private non-profit research centres, universities, and colleges, or promote resource sharing and collaboration among individuals at these institutions. They are referred to henceforth as academic and/or research networks. Exceptions among those included are networks that represent the only network in a country, or are important because they are interconnected with many other networks.

Location The geographic location of a network can range from metropolitan networks joining a number of universities within a single city, through regional networks that link institutions across an number of states or provinces, national networks that join hosts and other networks country-wide, to international networks that join institutions or networks in different countries. The geographic area each encompasses is outlined briefly.

Configuration This category describes network operating characteristics of the network including underlying protocols in use, services supported, speed of data transmission, and physical media used in constructing the links.

There are a great variety of protocols that underlie computer networks. Protocols such as TCP/IP, X.25, ISO-OSI, DECNET, Colored Books, UUCP, XNS, and SUN-III, are all used across the large number of networks. Although most network governing bodies have stated a commitment to migrate to OSI standards, the majority have yet to actively begin conversion. To complicate matters, protocols can also vary *within* a single network, upper layer protocols may differ from lower layer protocols, and two different protocols may run concurrently on one system. While networks running dissimilar protocols can be interconnected, such interconnection requires protocol conversion.

Networks also differ in the services they provide, an issue related to the underlying protocols in use. The most common services can include mail, file transfer, remote login, remote job entry, and conferencing. Each major protocol suite typically has its own implementations of the services; for example, file transfer is defined as the FTP application protocol within TCP/IP, and as FTAM within OSI. Not necessarily all services within a protocol suite, however, will be supported by a given network. Most networks support email and file transfer, but conferencing is supported by only a few. In the end, the types of things that users can do on a network are dependent upon the network's underlying protocol, and the services chosen for implementation.

The ways in which computer networks are constructed physically are diverse as well. The transmission media can be cable, fiber optics, satellite, radio, and microwave, or a mixture of any or all five. The links a network uses can be leased lines, dial-up links, high-speed fiber optic backbones or, as is common in Europe, commercial public packet-switching data networks (PDNs). Again, a single network may possess a mixture of links; for example, one network may have a high-speed backbone linking major centres, leased lines linking intermediate hosts, and dial-up access for individual users. Speed of data transmission is dependent on transmission medium, link type, and traffic flow. These can range from 300bps on dial-up lines, to 10Mbps on an Ethernet, and can range as high as 1Gbps.

Governance Governance involves the organization, administration, and maintenance of a network. For most networks, governance is provided by a centralized group comprised of representatives from member institutions.

Financing Networks can be government-funded or self-supporting. Government funding can take the form of on-going support or simply an initial start-up or seed grant. Self-supporting networks charge those who use the networks. This usually means annual fees to individual institutions rather than individual users. In some cases, institutions must pay for a link to another site, and then support the flow of communication through their node to other parts of the network.

Membership This category lists, when available, the institutions and organizations that participate in the network. When the list is too long, simple numerical counts are given.

Connections Most networks are interconnected with one or more other networks. These interconnections can become rather complicated, however, with direct links to some networks and indirect links to others through a third intermediary network. At times, a loose hierarchy is visible among interconnections. An example of this is NSFNET, a backbone stretching across the U.S. connecting a number of mid-level networks. The hierarchy becomes quickly tangled, however, as many mid-level networks are themselves connected independently to networks outside NSFNET. The result seems to be no consistent pattern among network interconnections. The service that is most commonly available across network boundaries is electronic mail. Connections such as remote login and other interactive services are less often achieved. This category lists the networks that are accessible from the particular network under discussion.

The following sections describe the research networks that currently in place around the world. Each section is divided by major geographic zone. If international networks are present for that zone, they are described at the beginning of each section.

A.1 North America

A.1.1 International networks in North America

a) The *Internet*

Name — The name of this network is derived from "internetwork" or "internet" meaning a collection of interconnected networks. This term spelled with a small "i" can be used generically to describe any such collection. Internet spelled with a capital "I", however, refers specifically to the *Internet*, a set of largely TCP/IP-based networks found in the United States, Canada, as well as some other continents.

Purpose — The *Internet*'s purpose is to facilitate the collaboration and the sharing of information among government agencies, educational institutions, and industry. It is the top-level metanetwork that connects other, also very large, networks which can contain component networks themselves. It is estimated to have anywhere between 40,000 and 500,000 hosts online. Its main backbone networks are NSFNET and MILNET (a military network not discussed in this paper).

Location — The networks that compose the *Internet* are located all across the U.S. and Canada. Some parts also extend to other countries such as Europe, Australia, and Mexico.

Configuration — Largely TCP/IP, although some individual networks run TCP/IP over X.25. The major application layer protocols supported are FTP, SMTP, and TELNET, for file transfer, mail and remote login, respectively. The links between constituent parts consist of a variety of technologies, for example, geosynchronous satellites (connecting the U.S. East and West coasts), leased lines, Ethernets, and dial-up lines. Because of the diverse nature of the component networks, no single statement of speed can be made; they range from 1200bps dial-up links to 1 Gbps or above (Lynch, 1991).

Governance — Administration is done by the Defence Data Network Information Center at SRI International in California and the Network Operations Center at Bolt, Beranek and Newman in Cambridge, Massachusetts.

Financing — The *Internet* is funded by grants from the U.S. government; campus-area networks are funded inhouse.

Membership The institutions that make up the *Internet*'s membership are too numerous to list because the *Internet* is a network of networks, each having its own member institutions. Many of the *Internet*'s constituent networks follow in the next section on the United States.

Connections The *Internet* has links to BITNET (US), JANET (U.K.), ARISTOTE (France), NORDUnet (Scandinavia), EUnet (Europe), JUNET (Japan), AARNET (Australia), CAnet and CDNnet (Canada), and ITESM (Mexico), among others. There is a planned connection to IXI (a European OSI backbone network).

A.1.2 The United States

a) NSFNET

Name National Science Foundation Network.

Date 1987

Purpose NSFNET is a large, general purpose research network that is one of the most well-developed in the U.S. It was originally created to connect supercomputer centres, but now provides other services. It has a number of specifically stated purposes. They are to:

- advance scientific collaboration on a national scale
- widen access to NSF funded supercomputer centres
- speed dissemination of research results
- enhance education through computer networking
- provide an experimental platform for network research
- establish leadership in research networking technology.

It has a three level structure consisting of:

- a transcontinental backbone that connects a number of mid-level networks
- mid-level networks consisting of regional, specific-discipline, and super-computing networks
- campus networks that connect to mid-level networks.

Location The NSFNET backbone extends throughout the continental United States. It has links to the U.K Europe, Canada, Mexico, Scandinavia, Japan, and Australia.

Configuration TCP/IP protocols are used on the backbone and most mid-level networks. Some component networks use DECNET. Services supported are remote login, file transfer, mail, and conferencing. By July, 1989, NSFNET had 19 physical links giving most nodes T1 links over primarily fiber optic cable. NSFNET is, as of Spring, 1991, mostly running at T3 speeds (Lynch, 1991). A fiber optic line also extends across the Atlantic that will link NSFNET with European networks at T1 speeds (NSF Flyer, 1989).

Governance Overall management was originally done by Merit, a consortium of eight Michigan Universities. As of late 1990, management is performed by a not-for-profit organization, Advanced Network and Services Inc. (ANS) created by ANS by Merit, IBM, and MCI to manage and operate the backbone.

Financing NSF provided 14 million dollars to Merit, IBM, and MCI Telecommunications to re-design and expand the existing NSFNET backbone. The State of Michigan Strategic Fund also contributed 5 million dollars toward the effort. The mid-level networks are funded independently, though some other funds come from NSF.

Membership The mid-level networks that compose NSFNET are BARRNet, JVNCNet, Merit, MIDnet, NCSAnet, NorthWestNet, NYSERNet, PSCnet, SDSCnet, Sesquinet, SURAnet, THEnet, USAN, and WESTNET. It was estimated that as of May, 1990, there were approximately 300 institutions connected (Huray and Nelson, 1990).

Connections NSFNET is part of the *Internet*. Outside connections include BITNET, CSNET, UNIDATA, and CDNnet through BCnet and Onet. Connections are in the offing for JANET, NORDUnet, and ISESM. There are test connections to ARISTOTE.

b) NSFNET mid-level networks

CERFnet

Name California Educational and Research Federation Network.

Date 1988

Purpose The purpose of CERFnet is to advance education and science by facilitating the interchange of information among research and educational institutions. It also provides supercomputer access, and will act as a testbed for network experimentation.

Location CERFnet is a regional network in southern California.

Configuration Protocols and services are TCP/IP. Link speeds are 9600bps to T1.

Governance Administration is done by a board of directors from CERF comprised of elected representatives from each member institution.

Financing NSF (National Science Foundation) startup grant, with partial funding thereafter for three years. Other monies will come from private industry and individual institutions.

Membership

- San Diego Supercomputer Center
- California State University Fresno
- University of California at Irvine
- California State University
- CO/SWRL Facility

In addition to those listed above, 38 other Californian centres for research or education have committed to join.

Connections CERFnet is part of the *Internet*

NYSERNet

Name New York State Educational and Research Network.

Date 1985

Purpose To support collaboration between universities, government agencies, and private business.

Location Primarily New York State.

Configuration TCP/IP-based. It has a T1 infrastructure and 9600bps and 56Kbps links to areas in neighbouring states. Network operators are currently working on increasing library access.

Governance Originally operated by NYSERNet, Inc., a consortium of component universities, NSF, the State of New York, and several private companies. Early in 1990, however, Performance Systems International Inc. (PSI), a provider of value-added network services, signed a five year contract with NYSERNet Inc. to take over the management of that portion of the *Internet* (Jackson, 1990). This is part of a "grass-roots" movement to commercialize the *Internet* proposed by a group of users who were dissatisfied with the management structure and costs of the present internetwork (Jackson, 1989).

Connections NYSERNet is connected to NSFNET, and BITNET. NYSERNet is a NSFNET regional network.

JVNCNet

Name John van Neumann Supercomputer Center Network.

Purpose JVNCNet is a consortium that connects several supercomputing centres.

Location Located mainly in the northeastern states of New Jersey, Pennsylvania, New York, Conneticut, Rhode Island, Massachusetts, and New Hampshire. There are also connections in Colorado and Arizona.

Configuration TCP/IP is used with IGRP and RIP (Routing Information Protocol) for internal routings, EGP (Exterior Gateway Protocol) for exterior routing. Nine T1 links arranged in a star around JVNC join the centre to other computing centres.

Governance Administered by several bodies according to function including Network Information Services, Network Operations, Network Installation and Maintenance, and Network Engineering. Responsibility for these bodies is divided among JVNCNet personnel and NSFNET personnel.

Membership Membership includes 22 colleges and universities located in the eastern U.S. It also includes the University of Colorado and the University of Arizona.

Connections BITNET, TYMNET, and NORDUnet.

SURAnet

Name Southeastern Universities Research Association Network.

Date 1987

Purpose SURAnet is designed to provide access to supercomputers at the University of Georgia, Florida State University, Virginia Tech, and the University of Alabama.

Location The southeastern U.S. states of Maryland, Virginia, West Virginia, Kentucky, Tennessee, Alabama, Mississippi, Louisiana, Florida, Georgia, South Carolina, North Carolina, and the District of Columbia.

Configuration TCP/IP-based.

Governance Managed by the principal authors of the original NSF proposal to build SURAnet.

Financing Provided by NSF.

Membership Membership includes 35 research universities in the southern U.S.

Connections SURAnet is an NSFNET mid-level network and is part of the *Internet.*

OARnet

Name Ohio Academic Resources Network.

Date 1987

Purpose It is a general purpose network that connects most academic institutions in Ohio. Its uses include collaborative research, library access, and supercomputer access.

Location Ohio.

Configuration The network supports TCP/IP, DECNET and NJE. The ultimate goal of the network is to use OSI protocols. TCP/IP has been identified as the migration path between DECNET and NJE, and OSI protocols. Links are 56Kbps with Proteon routers.

Governance Administered by the Ohio Supercomputer Center. Network policies are set by the Stateside Network Users' Council composed of a representative from each member institution.

Financing Funding is provided by NSF, and the Ohio state government, but most is provided by the Ohio Board of Regents.

Membership
- Battelle Memorial Institute
- NASA Lewis
- Ohio Supercomputer Center

Connections OARnet is now a NSF backbone link (Lynch, 1991).

CICNet

Name Committee on Institutional Cooperation Network.

Date 1988

Purpose CICNet connects institutions of higher education in the midwestern U.S.

Location Illinois, Iowa, Michigan, Minnesota, Ohio, and Wisconsin.

Configuration TCP/IP-based with T1 links.

Governance The network is incorporated as CICNet Inc. It has a board of directors consisting of a representative from each institution, and one from the CIC.

Financing A start up grant was received from NSF in 1987. Additional funding is provided by member institutions. The eventual goal is self-sufficiency.

Membership
- Indiana University
- Michigan State University
- North Western University
- Ohio State University
- University of Chicago
- University of Illinois at Urbana-Champaign
- University of Iowa
- University of Michigan
- University of Minnesota
- University of Wisconsin

Connections CICNet is an NSFNET mid-level network and is part of the *Internet.*

Merit

Name Merit is the name of a consortium of Michigan universities.

Date 1972

Purpose One of the oldest networks in the U.S., it began as an inter-university network and is now an NSFNET mid-level network.

Location Michigan.

Configuration The network is based on older pre-TCP/IP locally developed protocols, but IP has been implemented. X.25 services are also provided. Long distance lines are 9600bps, with some 56Kbps links.

Governance The overall network is managed by Merit Inc. Local parts are separately administered.

Financing Seed money was provided by NSF and by the member institutions. It is currently self-supporting.

Membership

- Central Michigan University
- Eastern Michigan University
- Michigan Technical University
- Oakland University
- University of Michigan
- Wayne State University
- Western Michigan University

Connections Merit is a mid-level NSFNET network and a component of the *Internet.* Other connections are to Telenet and TYMNET.

MIDnet

Name The name MIDnet is derived from the fact that the network is located in the American Midwest.

Date 1985

Purpose The network links 15 midwestern universities.

Location Iowa, Illinois, Missouri, Arkansas, Oklahoma, Kansas, and Nebraska.

Configuration TCP/IP-based, though some participating organizations support DECNET or SNA. Links consist of 56Kbps leased lines arranged in a ring encompassing the seven states.

Governance The network is governed by a board of directors composed of representatives from each member institution.

Financing NSF provided a three-year start up grant. The ultimate goal is self-sufficiency.

Membership

- Iowa State University
- Kansas State University
- Oklahoma State University
- Oklahoma University
- University of Arkansas
- University of Iowa
- University of Nebraska
- University of Tulsa
- Washington University

Connections MIDnet connects to the NSFNET backbone at the University of Nebraska, Lincoln, and is a component of the *Internet.*

NCSAnet

Name National Center for Supercomputing Applications Network.

Date 1986

Purpose NCSAnet connects NCSA supercomputers to the NSFNET backbone

Location Indiana, Illinois, and Wisconsin.

Configuration Based on the TCP/IP suite of protocols. Link speeds are T1, 65Kbps, and 14.4Kbps.

Governance The network is managed by NCSA in cooperation with network staff at the University of Illinois.

Financing Initially funded by NSF; the University of Illinois, Chicago, and NCSA shared the cost of a T1 line. Other institutions pay for their own 56Kbps lines.

Membership

- Argonne National Laboratory
- Bradley University (planned hook-up)
- Fermi National Accelerator Laboratory
- Illinois Institute of Technology
- Indiana University
- National Center for Supercomputing Applications
- Northwestern University
- Notre Dame
- Purdue University (planned hook-up)
- University of Chicago
- University of Wisconsin

Connections Connects to NSFNET as a mid-level network and a part of the *Internet.* Other connections are to MIDnet, and MRNet.

MRNet

Name Minnesota Regional Network.

Date 1987

Purpose MRNet is a NSFNET regional network connecting a number of higher education institutions, businesses, and supercomputer centres.

Location Minnesota.

Configuration The network uses TCP/IP protocols exclusively. The backbone network is an Ethernet. Speeds range from 9600bps to 56Kbps; it also has 10Mbps Ethernet links.

Governance MRNet is administered by the MRNet Organization, an informal group of network member representatives.

Financing Initial seed funds came from NSF, with member organization covering operating expenses. Members also paid for their own connections. Its objective is to be self-supporting by the end of 1989.

Membership

- 3m Corporation
- Carlton College
- Control Data Corporation
- Cray Research, Inc.
- ETA Systems
- Honeywell
- Mayo Foundation
- Minnesota Supercomputer Centre
- Network Systems Corporation
- St. Olaf College
- University of Minnesota

Connections MRNet is connected to NSFNET, BITNET, and USENET. It is also part of the *Internet*.

WESTNET

Name The name, WESTNET, is derived from the fact that the network extends into five states in the American west.

Purpose WESTNET is an academic and industrial network. It also connects many smaller regional networks such as the New Mexico Technet, a TCP/IP optical fiber net, and the Colorado Supernet, a 56Kbps TCP/IP network. It allows the smaller networks to be connected together directly without routing through the NSFNET backbone thereby reducing congestion.

Location Colorado, New Mexico, Arizona, Utah, and Wyoming.

Configuration It is a TCP/IP internetwork with 56Kbps long-haul links.

Financing A three-year NSF start-up grant is due to expire in 1991.

Membership WESTNET connects 16 academic institutions and four industrial sites.

Connections WESTNET is an NSFNET mid-level network. It also has connections to NorthWestNet and NYSERNET

NorthWestNet

Name The name stems from the networks location in the northwestern part of the U.S.

Date 1987

Purpose Provides network service to six northwestern U.S. states. After incorporation as a non-profit organization, the network will extend service to smaller universities, government labs, and commercial companies.

Location North Dakota, Montana, Idaho, Washington State, Oregon, and Alaska.

Configuration The network runs TCP/IP over Proteon routers. Links between campuses run from 9600bps to 56Kbps. There is a T1 digital radio link to NSFNET and a satellite link to Fairbanks, Alaska.

Governance NorthWestNet is managed by Boeing Computer Services, under contract to the Western Interstate Commission for Higher Education. There are plans to become a non-profit organization under the name of the Northwest Academic Computing Consortium.

Financing Financing for the network is fee-based, currently $5000 per annum for non-profit organizations, $25,000 for for-profit organizations.

Membership

- Boeing Computer Services
- Montana State University
- North Dakota State University (Fargo)
- Oregon Graduate Centre
- Oregon State University (Corvallis)
- University of Alaska
- University of Idaho (Moscow)
- University of Oregon (Eugene)
- University of Washington (Seattle)
- Washington State U (Pullman)

Connections NorthWestNet is a mid-level NSFNET network and a component of the *Internet.*

BARRNet

Name San Francisco Bay Area Regional Research Network.

Date 1986

Purpose BARRNet connects academic, private, and government research organizations in the San Francisco area. Growth in number of hosts and geographic area is expected.

Location Northern California.

Configuration The network uses TCP/IP with T1 links between campuses.

Governance Organized as an association of its participating and affiliated institutions.

Connections NSFNET, and SPAN. It is also part of the *Internet*.

THEnet

Name Texas Higher Education Network.

Date 1984

Purpose THEnet is a state network that connects academic, medical, research and corporate institutions.

Location Texas.

Configuration THEnet uses DECNET primarily, although TCP/IP is supported on some campuses. Links are 9600bps and 56Kbps using leased lines.

Governance THEnet is administered by the Office of Telecommunication Services of the University of Texas System.

Financing Financing is the responsibility of the member organizations.

Membership Membership includes over 40 universities, medical centres, government agencies, corporations, and research centres.

Connections NSFNET, BITNET, CSNET, and SPAN.

Sesquinet

Name Texas Sesquicentennial Network.

Date 1987

Purpose Connects research sites in Texas to each other and to the NSFNET backbone.

Location Limited primarily to the Houston area, although it is expected to expand state-wide.

Configuration TCP/IP-based, it uses 56Kbps leased lines with plans to upgrade to T1.

Governance Organization is provided by a steering committee at Rice University.

Financing Funded by NSF.

Membership
- Baylor College of Medical Technology
- Baylor College of Medicine
- Prairie View A & M
- Rice University
- Texas A & M
- Texas Southern University

Connections NSFNET.

USAN

Name University Satellite Network.

Date 1987

Purpose USAN connects its member institutions to the National Center for Atmospheric Research allowing supercomputer access.

Location Country-wide, although most sites are concentrated in the eastern half of the U.S.

Configuration The network uses KU-band satellite links that use Very Small Aperture Terminal (VSAT) technology. Speeds range anywhere from 9600bps over ground links to 224Kbps in the satellite broadcasts.

Membership
- Institute for Naval Oceanography
- Naval Research Laboratory
- Oregon State University
- University of Maryland
- University of Miami
- University of Michigan
- University of Wisconsin
- Woods Hole Oceanographic Institute

Connections There are links to BITNET and SPAN, NSFNET and the *Internet.*

e) MFEnet

Name Magnetic Fusion Energy Network.

Date 1975

Purpose Provides access to supercomputer facilities. Originally set up to connect physics departments doing nuclear fusion research, but has been expanded to include all U.S. Department of Energy projects.

Location The United States and Japan.

Configuration The network uses custom protocols called NSP over the IP protocol. There are plans to migrate to OSI.

Governance Administered by the Department of Energy. Day to day management is provided by National MFE Computer Center.

Financing Provided by the Department of Energy.

Membership 120 hosts in the continental U.S. and Japan.

Connections There are gateways to the *Internet.*

f) SPAN

Name Space Physics Analysis Network.

Date 1981

Purpose SPAN is a NASA general purpose network. It is designed to provided an infrastructure for space-related research.

Location Located largely in the U.S., with extensions to Canada, Japan, and Europe.

Configuration Based on DECNET Phase IV, it will eventually migrate to DECNET Phase V for OSI compatibility. Links between centres are 56Kbps.

Governance Administration is done by project managers, network managers, and routing centre managers.

Financing SPAN is funded by NASA.

Membership There are currently 2400 hosts on the SPAN network.

Connections There are gateways to the *Internet* and BITNET. JANET, FUNET, SUNET, and INFNET are also accessible.

A.1.2.1 Other important U.S. networks

a) BITNET

Name Because It's Time Network.

Date 1981

Purpose The network was originally created to allow collaboration among systems programmers at campus computing centres. It was established originally as a free communication network. It is not part of the *Internet*, although mutual access is available.

Location The hosts are concentrated mainly in the U.S., but nodes extend outside the States to Canada, Mexico, Argentina, and Asia.

Configuration It is a "store and forward" network, that is, a local host computer sends a given message to the next host on the path to the destination (Landweber and Jennings, 1986). The Network Job Entry (NJE) protocol is used with most hosts using the RSCS implementation on VM, or JES/NJE on MVS. Services provided are mail, file transfer, and interactive messages.

Governance BITNET is governed by the BITNET Information Center.

Financing Supported by membership fees.

Membership There are 397 sites in the U.S., and many more worldwide.

Connections NetNorth, GulfNet, CSNET, and indirect access to the *Internet*.

b) USENET

Name Users' Network.

Date 1979

Purpose USENET is one of the oldest and largest networks in the world. It originates in the U.S. but has since extended around the world. It is a news network rather than a mail or research network. It is mentioned because it has served as a model for other networks, such as its European equivalent, EUnet.

Location Five continents world-wide, with highest concentrations in the U.S.

Configuration USENET uses a flooding algorithm in which each host sends all news articles it receives to all other hosts it can reach online.

Governance The system is a cooperative network with a completely decentralized administration.

Financing There is no funding mechanism. Each host pays for its own connection.

Membership Membership is open to whomever has a UNIX machine and sufficient funds to make a connection.

Connections It does not have a mail service, so it is not connected to other networks in the usual sense. Some other networks do, however, support distribution of USENET mail.

c) UUCP

Name The name is derived from the UUCP protocol, meaning UNIX to UNIX Copy Program.

Date 1987

Purpose The network connects machines running the UNIX operating system. The UUCP protocol developed for this network is used quite widely in international connections.

Location Beginning in the U.S., it has since spread throughout the world.

Configuration UUCP protocol running over RS-232-C dial-up links. In Europe, it is sometimes used over X.25. There are also adaptations for its use over TCP/IP. Mail is the only service provided worldwide.

Governance This is a decentralized network with no political entity at its head.

Membership All UNIX-based systems that wish to apply.

d) HEPnet

Name High Energy Physics Network.

Purpose This network originated in the U.S. and has since expanded to other continents. Its purpose it to facilitate communication between researcher in this area.

Location HEPnet is both a major U.S. and European network.

Configuration The system is based on DECNET both because it can handle large files and can run over a variety of network protocols such as X.25 and Ethernet. Links are leased lines running at 9600bps and 56Kbps.

Membership Core high energy physics centres in the U.S. are:

- Argonne National Laboratory
- Brookhaven National Laboratory
- Fermi National Accelerator Laboratory
- Lawrence Berkeley Laboratory
- Stanford Linear Accelerator Center

Connections The *Internet* and BITNET.

A.1.3 Canada

a) NetNorth

Name Derived from, presumably, the fact that Canada is the northern part of North America.

Date 1984

Purpose NetNorth is an academic and research network. The main goals of NetNorth are to:

- enable electronic communications and foster cooperation between Canadian academic institutions and associated research institutions in both the private and public sectors

- provide connectivity to other networks and organizations of interest to the NetNorth consortium's members .

Location Canada.

Configuration NetNorth is BITNET in Canada with the same NJE protocols and services. Also supported are TCP/IP and DECNET. Link speeds range from 2400bps to 9600bps.

Governance The network is run by the directors of the NetNorth Consortium.

Financing Line costs are divided equally among participating institutions. Funds were also provided by a three-year commitment from IBM Canada in 1984.

Membership Membership is open to universities, colleges, non-profit research institutions, educational institutions, federal organizations, and provincial organizations. There are 59 member institutions Canada-wide.

Connections There is a direct connection to BITNET and an indirect connection to EARN. Connections are in the offing to NSFNET and the *Internet.*

(NetNorth 1989 Strategic Plan)

b) CDNnet

Name Canadian Network.

Date 1986

Purpose The purpose of CDNnet is to provide network access the research, educational, and advanced development agencies. It was the first network to actively use the CCITT X.400 messaging protocols.

Location Canada.

Configuration The network uses the Ean implementation of X.400 developed at the University of British Columbia. The network layer is X.25 and DECNET, with some parts utilizing TCP/IP. Long-haul links are X.25 at 2400bps, with a 19.2Kbps leased line to the NSFNET backbone. Services include X.400 mail and remote login (via Triple-X to hosts with X.25 service). FAX service is under development, as well as an X.500 Directory service.

Governance The network is administered at the CDNnet Headquarters at the University of British Columbia. Policy decisions are made by a committee consisting of representatives from each member organization.

Financing Based on membership dues and a grant from the National Science Educational and Research Council that was used to develop the X.400 implementation.

Membership There are 47 sites in Canada.

Connections CSNET, BITNET, NSFNET (and the rest of the *Internet*), and EUnet.
(CDNnet Reports, 1989)

c) CAnet

Name The name stands for Canada Network. It was also known briefly as NRCnet, after the National Research Council, its funding agent.

Date 1989

Purpose The network is based on the NSFNET model in the U.S. Similar to NSFNET, it provides a high-speed backbone linking mid-level networks in support of collabratory activity between academia, government, and industry. It is also intended to stimulate the transfer of technology between the public and private sectors.

Location CAnet will link networks in each of 10 regions in Canada.

Configuration TCP/IP is used to connect the regional networks at speeds of 56Kbps. The network will support T1 speeds and the line become affordable. Services include database access, supercomputer access, electronic mail, and file transfer. OSI features will be implemented as they are developed.

Governance Managed by CAnet Networking, Inc. consisting of a board of directors derived from the regional networks. It is operated by the University of Toronto Computing Services Department.

Financing Seed money was provided by the NRC. Self-sufficiency is expected after the five-year grant expires; half of all costs will be divided equally among the regions, the other half will be divided apportionally by expected usage.

Membership Mid-level networks joined by CAnet will be NLnet, NSTN, PEI Network, AccessNB, RISQ, Onet, MBnet, Sasknet, ARnet, and BCnet, as each is implemented.

Connections NSFNET and the rest of the *Internet*, NETNORTH, and CDNnet. (Siciunas, 1990).

d) CAnet mid-level networks

NLnet

Name Newfoundland and Labrador Network.

Date 1990

Purpose A provincial data communications network designed to support research, education, and technology transfer between member institutions. NLnet will promote the development of local databases and facilitate resource sharing.

Location Newfoundland.

Configuration Exclusively TCP/IP. A full range of network services will be provided over time.

Governance Exists as an association of its member organizations.

Financing NLnet is currently self-supporting.

Membership
- Cabot Institute of Applied Arts and Technology
- Department of Fisheries and Oceans
- Forestry Canada, Newfoundland and Labrador Region
- Institute for Marine Dynamics
- Institute for Fisheries and Marine Technology
- Memorial University of Newfoundland

Connections NLnet is a CAnet mid-level network.

NSTN

Name Nova Scotia Technology Network.

Date July, 1990

Purpose To allow universities, colleges, research institutions and private industry in N.S. to communicate with the rest of Canada.

Location Nova Scotia.

Configuration It is a TCP/IP backbone with three types of connections offered: 56Kbps with a router at client sites, 4800bps SLIP service (which allows TCP/IP over serial lines), and dial-up service.

Governance NSTN is a wholly owned, for-profit subsidiary of Software Kinetics Ltd. Operations services are contracted to Dalhousie University Computing Services.

Financing The Canadian government provided a 1.95 million dollar start-up grant. The network is expected to be self-supporting by 1991.

Membership

- Acadia University
- Bedford Institute of Oceanography
- Dalhousie University
- Mount Saint Vincent University
- NSTN Inc.
- Saint Francis Xavier University
- Saint Mary's University
- Software Kinetics, Inc.
- Whitman Benn

Connections NSTN is part of CAnet.

AccessNB

Name AccessNB is the New Brunswick/Prince Edward Island Educational Computer Network.

Date Formed in the late 1960's.

Purpose The network currently data communication and information transfer among N.B. and P.E.I universities. In the future, expansion will allow it to serve industry and research centres as well.

Location New Brunswick, Prince Edward Island.

Configuration The current network provides computer to computer links over leased lines running at 19.2Kbps. Services offered are RJE, mail, and ASCII and 3270 terminal service. They are currently testing TCP/IP. In its AccessNB incarnation, it will provide TCP/IP service to regional industries, government laboratories, and universities.

Governance Administration is done by a non-profit organization called CADMI/MTC. Operation is carried out by the University of New Brunswick's Computing Services Department.

Financing Currently funded by a government grant. In the future it hopes to be self-supporting from membership fees.

Membership

- Mount Allison
- St. Thomas University
- The University of Moncton
- The University of New Brunswick
- The University of Prince Edward Island

Connections AccessNB is part of CAnet.

The Prince Edward Island Network

The University of P.E.I. has undertaken two projects to study the feasibility of creating a network specific to Prince Edward Island. The first project will be to install a CAnet node at the University and provide access to a limited number of users including Agriculture Canada, the PEI government, and Holland College. The second project will determine the operational requirements of a PEI network.

RISQ

Name *Réseau interordinateur scientifique québécois.*

Date 1989

Purpose To support the educational and research aims of all Quebec universities in all areas. Public and private research laboratories may join if their use is compatible with RISQ objectives.

Location Quebec.

Configuration CISCO to CISCO connections at 9600bps, 19.2Kbps, 56Kbps, or 10Mbps

Governance RISQ is administered by an association of its members.

Financing Major funding comes from the *Centre de recherche informatique de Montréal* (CRIM). Other funding is provided by members.

Membership

- Atmospheric Environment Service (Dorval)
- Bell Canada
- Centre de recherche informatique de Montréal (CRIM)
- Concordia University
- École des hautes études commerciales (HÉC)
- École Polytechnique
- HEPnet (Canada)
- INRS-Telecommunication
- Institut de recherche de l'Hydro-Québec
- McGill University
- Université de Montréal
- Université de Sherbrooke
- Université du Québec
- Université du Québec à Montréal
- Université Laval

Membership is expected to expand to 12 universities, 5 research centres, and 5 government departments in 1991.

Connections Onet, NSTN, the *Internet* via JVNCNet, and CAnet.

Onet

Name Ontario Regional Network.

Date 1988

Purpose To facilitate communications between member institutions and to support their educational and research missions. Also provides access to a Cray supercomputer located at the University of Toronto.

Location Ontario.

Configuration Protocols in use are TCP/IP, NJE over TCP/IP, and some strictly NJE. DECNET is supported at Toronto. It uses 19.2Kbps or 56Kbps leased lines. The network supports email based on the Simple Mail Transfer Protocol and remote login based on TELNET.

Governance Onet is run by a non-profit association. It is managed by three committees. A Management Committee consisting of a representative of each member institution, a Technical Committee responsible for technical planning, and an Operations Committee, responsible for daily management.

Financing Fully funded by membership fees. Onet is currently moving toward a usage-based charging structure.

Membership Membership consists of a total of 36 institutions: 16 universities, 5 colleges, as well as several research centres, government departments, and private sector organizations.

Connections NSFNET, CAnet, and the *Internet*.

(Marshall, 1989)

MBnet

Name Manitoba Network.

Date June, 1990.

Purpose The network was set up to serve post-secondary institutions and other organizations involved in research.

Location Manitoba.

Configuration TCP/IP-based with line speeds of 56Kbps.

Governance Governed by a management committee consisting of representatives from the University of Manitoba, the University of Winnipeg, and the National Research Council (NRC).

Financing Start-up funds are provided by NRC. Each member pays for the line to the nearest node. Operating costs are met by members.

Membership In addition to the University of Manitoba, the University of Winnipeg, and NRC, membership will consist most typically of federal and provincial research laboratories.

Connections CAnet.

Sask#net

Specialists in Saskatchewan plan to create a network to provide services to universities, government agencies, and public and private research groups. Supercomputer access is a special interest. The network will serve as the CAnet node in this province. Initial funding will be provided by two universities in the province, and is intended to be self-supporting. Current members are the University of Saskatchewan and the University of Regina.

ARnet

Name Alberta Research Network.

Date January, 1990.

Purpose ARnet is an academic and research network, with plans to include members from the private and public sectors.

Location Alberta.

Configuration A TCP/IP network with 56Kbps links.

Governance The network is managed by the three founding members and operates on a non-profit basis.

Financing All operating funds are provided by the current members. Funding continues to be a problem.

Membership

Current membership consists of:

- University of Alberta
- University of Calgary
- Alberta Research Council

Membership is expected to expand to 10 educational and private institutions by the end of 1990.

Connections A link to the *Internet* has been established through the University of British Columbia.

BCnet

Name The British Columbia Regional Network.

Date 1987

Purpose To support education, research and development, and technology transfer.

Location British Columbia.

Configuration TCP/IP, DECNET, and X.25 are all used. Transmission media are a mixture of microwave, fiber optics and coaxial cable. All link speeds between 2400bps to 2Mbps are used. Services include remote terminal access, file transfer, electronic mail, and free access to file archives.

Governance BCnet exists as non-profit consortium of cooperating members. Management is provided by a board consisting of member representatives.

Financing Start-up funds came from federal and provincial grants. Members pay for equipment and communication links as well as pay annual dues to cover operating costs.

Membership

- Advance Light Imaging
- Consumers Software
- Consumers Software
- Dominion Astrophysical Observatory
- Health Care Systems, Richmond
- Institute of Ocean Science
- Macdonald Dettwiler and Associates
- Microplex systems
- Microtel Pacific Research
- Pacific Forestry Centre, Forestry Canada
- Simon Fraser University
- Softwords, a division of Press Porcpic Ltd.
- TRIUMF Particle Accelerator Facility
- University of British Columbia
- University of Victoria
- Wimsey Associates

Future members will include:

- British Columbia Institute of Technology
- Camosun College, Victoria
- Royal Roads Military College, Victoria
- Vancouver General Hospital, Vancouver

Connections CAnet, NSFNET.

e) ISTC Network

A feasibility study was commissioned in early 1990 by *Industry, Science, and Technology Canada* to investigate the possibility of creating a high-capacity, high-speed backbone network in Canada. This high-speed national communications network, also known as ISnet (Industry, Science Network), would provide greater functionality to the Canadian research, development, education, and information communities (Hickling, 1990). This network would not grow out of CAnet, discussed above, but would be a separate initiative along the lines of the plans for the National Research and Educations Network (NREN) in the U.S. (see Chapter 6 for further details about NREN). Some general characteristics envisioned for this network include:

- the provision of user friendly services to organizations involved in research, development, and education so that they can communicate, access special equipment, and access information resources

- the provision of an environment for the development and testing of next generation information technology equipment and services

- the creation of a national backbone network connecting regional and international networks at an initial capacity of 1.5Mbps to be increased in the future as require to keep the network at the forefront of technology

- the conformance to ISO standards as they evolve, as well as enable co-existence with TCP/IP and other protocols.

Benefits of this network are considered to include:

- improved international competitiveness for industry
- improved quality and productivity of research
- more efficient use of scarce resources
- greater opportunities for small or geographically remote organizations
- improvements in educational techniques.

(Hickling, 1990:v).

A.2 Europe

A.2.1 International networks in Europe

a) IXI

Name Pilot International X.25 Infrastructure Backbone Service

Date June 8, 1990

Purpose IXI is a pilot network backbone designed to connect research networks in Western Europe eliminating the constraints and restrictions encountered with connections using public international telecommunications service. The pilot project is the first major activity of the EUREKA COSINE project, lasting 12 months. The long term goal of IXI is to provide high-capacity, cost-effective communications between all European network users.

Location The network spans the whole of Western Europe including Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden (providing gateways to Finland, Iceland, and Norway), Switzerland, the U.K and Yugoslavia.

Configuration IXI is an X.25 backbone network running at a speed of 64Kbps with a possibility of an upgrade to 2Mbps. Services supported are electronic mail, file transfer, as well as directory and information facilities.

Governance The network is administered jointly by the COSINE Policy Group, representing CEC national governments, and RARE (*Réseaux Associés pour la Recherche Européenne*), representing the actual and potential network user community. PTT Telecom will be responsible for creating, operating, managing the internetwork in cooperation with other European PTTs. While governance of the IXI internetwork will be provided by COSINE for the short term, there are uncertainties about governance and funding beyond 1992 when the COSINE Project ends.

Financing The network is funded by the Commission of the European Communities (CEC) and other participating national governments. From 1991, the national networks will increase their share of funding (Stone, 1991).

Membership Membership is restricted to the researchers, including industrial research departments, and government agencies engaged in research.

Connections There are connections to major European research networks, many of which have a connection to North America.

(IES News, 1990a).

b) EUnet

Name The network of EUUG, the European Unix User Group.

Purpose EUnet is a European-wide cooperative news and mail network used for research and development. It provides users with modern communications facilities and interconnections with other networks. The network began as an extension of the USENET and UUCP systems in North America.

Location Western Europe.

Configuration Ordinary UUCP dial-up links are the most common for local links in EUnet. Connections that extend outside Europe are UUCP over X.25. Links between backbones are currently being changed to leased lines. In some places, the SLIP protocol (allows TCP/IP over serial lines) may take the place of X.25 because it is faster. TCP/IP is also sometimes used over X.25 with RFC877 encapsulation. UUCP application protocols are widely utilized for news and mail services, as in USENET. On the TCP/IP parts, NNTP is used for news transfer and SMTP for mail. The speed of most X.25 links is about 4800bps. There are detailed plans for migration to OSI protocols, however, EUnet users have expressed no desire to do so at this time.

Governance Each backbone host organizes communications within its own country. Meetings are held annually where administrators of the backbone hosts decide EUnet strategies. The European UNIX systems Users Group acts, when necessary, as EUnet's voice.

Financing Funding is provided by individual hosts and users. All costs are shared equally among the members. Intercontinental links, however, are the responsibility of individual message senders.

Membership There is one EUnet backbone host in each member country in Europe. There are hosts in Austria, Belgium, Denmark, Finland, France, Germany, Greece Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the U.K and Yugoslavia.

Connections EARN, JANET, DFN, JUNET, SDN, AARNET, RangKoM, CSNET and UUCP.

c) EARN

Name European Academic Research Network.

Date 1983

Purpose EARN is a backbone network connecting national and campus networks into a continental network. Most sites are universities and government research facilities. Design of the network was based on BITNET in North America. Discussions have taken place between the three largest European networks (EARN, EUnet, and HEPnet) about the potential of a shared infrastructure.

Location There are nodes in all Western European countries, as well as Austria, Yugoslavia Cyprus, Turkey, Israel, Algeria, Ivory Coast, Morocco, Tunisia, and Egypt.

Configuration Based on the same technology as BITNET, the underlying protocols are NJE. There are plans to migrate to OSI protocols. Initially, this process will take the form of running NJE protocols over OSI to preserve the current applications. Higher level OSI protocols such as X.400 will be implemented later.

Governance Administered by a board of directors based in France consisting of elected representatives from member countries.

Financing Finances are the responsibility of each EARN node. Each must establish an international link with another member. Costs of links outside Europe are born by all members as a percentage of Gross National Product.

Membership There are currently more than 600 hosts at more than 500 institutions.

Connections There are gateways to JANET, HEANET, EUnet, and ILAN. Mail can be sent to BITNET and the *Internet*.

A.2.2 United Kingdom

a) JANET

Name Joint Academic Network.

Date 1984

Purpose JANET was created to provide a managed backbone network between academic and research institutions, and to provide access to the public network and networks outside the U.K. JANET is currently undergoing a major upgrade that will provide greater performance and will be known as JANET MK II (Cooper, 1990).

Location The United Kingdom and Northern Ireland.

Configuration Local networks connected to JANET are X.25 campus switches, Cambridge Rings, or Ethernets. The long-haul network layer is X.25, while applications at the higher layers are based on the Coloured Book suite of protocols. Links are leased lines running at 2Mbps. Site connections are increasingly at 2Mbps, rising from 64Kbps (of which some are multiple lines). Services supported are mail, file transfer, remote login, job transfer and manipulation. During 1991, JANET plans to migrate to basing its electronic mail on ISO 10021/X.400 (88) and well as initiating X.500-based Directory services (Gillman, 1990).

Governance The primary management body is the Network Advisory Committee (NAC) which It draws its membership from the Computer Board, and various Research Councils, universities, polytechnics, and the National Users Group. (Stone, 1990c). Technical and administrative support is done by the Joint Network Team (JNT), bases that Rutherford Appleton Laboratory, Didcot, Oxfordshire. In 1991, it is intended to restructure the management of JANET within the proposed UK networking Association (Stone, 1991).

Financing JANET is largely funded by the Computer Board. No direct charges are made to academic institutions or their users.

Membership

As of December, 1990, over 2000 hosts were connected (Stone, 1991). These include:

- universities, polytechnics, and some other colleges
- laboratories or institutes supported by Research Councils and government departments
- other institutions and organizations involved in collaborative research, including the industrial research community.

Other organizations are connected on a fee basis (e.g. SLS Information Systems and OCLC).

Connections Other European research networks via IXI, the U.S. *Internet* via NSFNET, public and research X.400 networks, EARN, EUnet, and SPAN (Stone, 1991).

b) UKnet

Name UKnet is the British portion of EUnet.

Location Britain.

Configuration Depending upon individual sites, protocols in use are Coloured book (for sites that are standard JANET sites) and UUCP. Services are those associated with UUCP. There are plans to participate in the EUnet migration to OSI.

Governance The network is run from the University of Kent at Canterbury in cooperation with the United Kingdom UNIX systems Users' Group.

Financing Mainly derived from charges to user sites, although there is some government funding.

Membership There are 300 registered sites.

Connections Those associated with EUnet.

A.2.3 Ireland

a) HEANET

Name Higher Education Authority Network.

Date 1985

Purpose The purpose of the HEANET is to provide researchers and educators access to hardware and software and to promote information exchange. It also provides connection to other networks outside Ireland.

Location Ireland.

Configuration Lower layer protocols are X.25 because the network uses Ireland's domestic X.25 packet-switching network, EIRPAC. High level applications are based on Coloured Book protocols. Application protocols are Green Book and Blue Book, for mail and file transfer, respectively. Remote login is supported by X.25 PADs.

Governance HEANET is administered and funded by the Irish Higher Education Authority.

Membership The network currently connects seven colleges throughout Ireland.

Connections There are gateways to EARN, the *Internet*, and EUnet.

A.2.4 Scandinavia

a) NORDUnet

Name Nordic University Network.

Date 1986

Purpose NORDUnet is an international network connecting the Nordic countries created to provide network services for research and development, and to establish good inter-Nordic network relations.

Location Sweden, Denmark, Norway, and Finland.

Configuration The backbone will support X.25, OSI, CLNS, TCP/IP, DECNET, and NJE in order to connect with networks that already use those protocols. Application layer protocols are Triple X, X.400, and FTAM. 64Kbps leased lines are used in the connections.

Governance The network is administered by NORDUnet itself.

Financing Funding is supplied by the Nordic Council of Ministers.

Membership As of 1989, there were 2000 to 3000 hosts online.

Connections SUNET, DENet, UNINETT, FUNET, EARN, EUnet, HEPnet, and NSFNET.

b) UNINETT

Name University Network.

Date 1978

Purpose The intention of UNINETT's management is to eventually connect all higher educational institutions in Norway.

Location Norway.

Configuration Inhouse UNINETT protocols are used, as well as DECNET, and Coloured Book protocols. Services provided are remote login, file transfer, conferencing, mail, and a directory service.

Governance The Norwegian Ministry of Cultural and Scientific Affairs has set up an organization to govern UNINETT.

Financing Funding is provided by the Norwegian Science Research Council and the participating institutions.

Membership There were 11 nodes as of 1987.

Connections EUnet, EARN, HEPnet, and the *Internet*. UNINETT is a component of NORDUnet.

c) FUNET

Name Finnish University Network.

Date 1984

Purpose The purpose of the network is to provide network service to Finnish universities and research centres.

Location Finland.

Configuration Most hosts are TCP/IP-based, while a few others use DECNET and NJE. Applications supported are mail, conferencing, file transfer, remote job entry, remote login and interactive graphics. OSI application protocols are tested at some sites including Ean, MRX.400, MHS, VM, and ISODE. Leased lines are used at speeds varying from 14.4Kbps to 64Kbps.

Governance Administered by the FUNET project.

Financing Non-academic organizations pay for the cost of their connections. Dues are collected from universities to cover the costs of national links.

Membership Most sites are universities, however, an increasing number of private companies and government agencies are connecting to the network. Nevertheless, use of the backbone is restricted to endeavors which further research and education. It currently connects 16 universities and more than 20 research institutes.

Connections Direct connections to EARN, EUnet, and CSNET. NORDUnet, of which FUNET is a part, will soon provide the connections to these networks as well as to NSFNET, HEPnet, and SPAN.

(Soini, 1989)

d) DENet

Name Danish Ethernet Network.

Purpose Provides nationwide access to computers at the Danish Computing Centre and those under the control of the Ministry of Education.

Location Denmark.

Configuration Both TCP/IP and DECNET are used. Ethernets are connected with MAC level bridges over 64Kbps or 126Kbps leased lines.

Membership The network connection department at Danish universities with local Ethernets.

Connections DENet is part of NORDUnet.

e) SUNET

Name Swedish University Network.

Date 1980

Purpose The network connects local and regional universities in Sweden. Its purpose is to provide data communications that will benefit the universities.

Location Sweden.

Configuration Local Ethernets are connected by 64Kbps lines to a national Ethernet. TCP/IP, DECNET, and NJE EARN logical networks are all carried over the SUNET backbone Ethernet. It is also possible to carry OSI protocols over SUNET, and full migration to OSI protocols is planned. There are currently X.400 message handling and X.500 directory pilot projects underway. Other services available over SUNET include file transfer, supercomputer access, a fax gateway, and a few information services.

Governance Operation and management of the network is spread around the different universities, although it is coordinated the Umea University Computing Centre.

Financing Early stages of SUNET development were financed by the Swedish Science Foundation. Ongoing funds come from the *Universitetsoch Högskoleämbetet.*

Membership

- The Medical University of Stockholm
- The Swedish University of Agriculture
- The Technical University of Gothenburg
- The Technical University of Stockholm
- University of Gothenburg
- University of Linkoping
- University of Lulea
- University of Stockholm
- University of Umeå

- University of Uppsala
- University of Lund

Connections FUNET is part of NORDUnet. Other connections include SPAN, HEPnet, the *Internet*, EARN, and EUnet.

(Wallberg, 1990)

f) EUnet in Iceland

Configuration There are UUCP links over leased and dial-up lines and X.25 links running from 1200bps to 9600bps. A connection to NORDUnet using TCP/IP over X.25 was expected at the end of 1989.

Governance EUnet in Iceland is administered by the Icelandic UNIX systems Users' Group.

Financing Sites are charged on a usage basis. Other monies come from the Icelandic Marine Research Institute.

Membership There are 20 registered EUnet hosts in Iceland.

Connections NORDUnet.

A.2.5 Germany (Formerly West Germany)

a) DFN (WIN)

Name *Deutsches Foerschungsnetz*. It is also known as WIN or *Wissenschaftnetz* (Science Network) (IES News, 1990b).

Date Around 1984.

Purpose DFN is a national research network that connects every university, college, and research laboratory in former West Germany. It is designed to develop protocols and implementations in the OSI suite of protocols, and to facilitate resource sharing and collaboration among researchers, both nationally and internationally.

Location The territory that was formerly West Germany, although connections are currently being planned in former East Germany (see A.3.2)

Configuration Network layer is X.25. File transfer and remote job entry application protocols are of an inhouse design, but are OSI compatible. X.400 mail is utilized. DFN is very involved in the implementation of OSI protocols particularly TP0, FTAM, and X.400 (Popper, 1990).

Governance The German Ministry of Research and Technology has a team of personnel working on the DFN project, however, overall administration is done by DFN-Verein in Berlin.

Financing Funded by the Ministry of Research and Technology.

Membership There were 65 hosts online as of the end of 1987. The DFN association itself is comprised of approximately 150 members from universities, research laboratories, and industry (IES News, 1990b).

Connections Gateways exist to EUnet, EARN, and CSNET.

b) AGFnet

Name AGFnet stands for *Arbeitsgemeinschaft der Grossforschungseinrichtungen* (AGF); the Association of National Research Centers of the Federal Republic of Germany.

Date 1988

Purpose AGFnet connects AGF research centres and German universities.

Location West Germany.

Configuration The network layer is X.25 using 65Kbps leased lines. It also supports TCP/IP and DECNET over the same physical links. AGFnet provides EARN with NJE service. OSI applications such as X.400 MHS and Triple-X are supported.

Governance AGFnet is both managed and funded by AGF.

Membership There are currently 12 organizations connected to the backbone each with its own network.

Connections EARN.

c) Dnet

Name Dnet is the German branch of EUnet.

Date 1983

Configuration Basic transport protocol is UUCP, while internal German links are X.25 PDN.

Governance Dnet is managed by a team of students and one staff member at the *Universität Dortmund.*

Financing Many facilities are provided by *Universität Dortmund*, however, link costs to EUnet are member-paid with an academic discount.

Membership As of November 1988, 146 companies and universities were connected.

Connections Connections are to CSNET, DFN, and EARN.

A.2.6 France

a) REUNIR

Name *Réseau des Universités et de la Recherche*, or Network of Universities and Research.

Purpose REUNIR is a metanetwork that connects many French universities and research institutions for the purpose of supporting research.

Location France.

Configuration Component networks consist of private X.25 networks that interconnect over *TRANSPAC*, the French PDN, SNA networks connecting mainframes, and local area networks running IP over X.25. Connections are either leased line or Telecom-1 X.21 satellite service. Speeds range from 4800bps to 2Mbps. Services provided are mail, file transfer, and remote login. There are plans to migrate eventually to OSI protocols.

Governance Administration is performed by two directors assisted by host administrators and a technical team.

Financing Monies come from the National Ministry of Education and *Centre National de la Recherche Scientifique* (CNRS).

Membership Basic membership consists of the national universities and CNRS. There are 10 large sites and 50 smaller ones.

Connections EARN, Fnet, and EUnet.

b) ARISTOTE

Name *Association de Réseaux Informatique en Système Totalement Ouvert et Tres Elaboré*, or the Association of Information Networks in a Completely Open and Very Elaborate System.

Date 1987

Purpose The network represents an association of French research institutions that share the goal of developing networking technology.

Location France.

Configuration Some members' internal networks are TCP/IP, but communication between them is done with X.25. International transmission is accomplished with X.25 over *TRANSPAC* at 4800bps to 9600bps. Faster connections are planned utilizing the French PTT's TELECOM-1 X.21 satellite and ISDN network, *numeris*. Services include remote login, X.400 mail, and a directory service called THORN.

Governance The network is structured as a non-profit organization with a Committee of Directors and an Assembly of Members. It also has technical working groups to handle specific areas.

Membership

- National Space Studies Centre
- Atomic Energy Commission
- National Telecommunications Centre
- Electricité de France
- National Research Institute for Computer Science and Automation
- National Institute of Nuclear Physics and Particle Physics
- Vector Calculus Centre for Research
- School of Mines of Saint-Etienne
- Laboratory of Automation and Systems Analysis or the National Centre of Scientific Research

Connections Fnet, CSNET, and the *Internet*.

c) FNET

Name Fnet is the French branch of EUnet.

Date 1983

Location France.

Configuration The most common protocol in use is UUCP, although Fnet contains one small TCP/IP network. Most links are carried over *TRANSPAC*, the French PDN, at 4800bps. There are plans to migrate to OSI.

Governance The system in managed by the *Institut National de Recherche en Informatique et Automatique* (INRIA) in cooperation with the French UNIX systems Users' Group (AFUU).

Financing The majority of funds come from INRIA with funds also contributed by AFUU. A user fee schedule has been proposed.

Membership Membership is composed of 123 hosts that include public research organizations (24%), universities (23.5%), computer manufacturers (19%), software vendors (15%), user groups (11%), and private research organizations (5%).

Connections ARISTOTE, REUNIR, EARN, and CSNET. A discussion is underway to connect Fnet to the *Internet*.

A.2.7 The Netherlands

a) SURFnet

Name *Samenwerkingsorgaisatie computerdienstverlening in het Hoger Onderwijs en Onderzoek*, or Cooperational Organization for Computer Services in Higher Educational Research.

Date 1989

Purpose SURFnet is a research and higher educational network and is part of EARN.

Location The Netherlands.

Configuration Part of the network uses DECNET and part uses NJE. Most links are 9600bps. A 64Kbps backbone using COSINE X.25 has been operational since January, 1990 The RFC822 mail format is implemented as a step on the migration path to X.400, although several hosts are already using this messaging protocol. Services include electronic mail, file transfer, remote login, remote job entry, conferencing, and bulletin boards.

Governance Management of the physical network is done by the Dutch PTT. General administration, user support, and protocol migration is the jurisdiction of SURFnet B.V. (Inc.) (Quarterman, 1990).

Financing Financing is provided by Foundation SURF, a public company 49 percent of which is owned by the Dutch PTT.

Membership There are 85 organizations currently connected.

Connections There are gateways to EARN, BITNET, the *Internet*, and EUnet. (Buitelaar, 1990)

A.2.8 Switzerland

a) SWITCH

Switch is an X.25 research network in Switzerland. No other information is available.

b) CERN

The Network of the European Laboratory for Particle Physics, or CERN, is a worldwide physics network based in Switzerland. It is connected to every European continental network. Due to the large number of machines connected, protocols that must be accommodated are TCP/IP, DECNET, SNA, TP4, and X.400.

A.2.9 Austria

a) ACONet

Name *Akademisches Computer Netz*, or Academic Computer Network.

Date 1982

Purpose ACONet is a long-haul research network that connects most research or higher educational institutions in Austria.

Location Austria.

Configuration It uses *Datenvermittlungssystem Nordrheinh-Westfalen* (DVS-NW) protocols. Services supported are file transfer, remote login, and remote job entry. It joins campus LANs using Austria's X.25 PDN, *DATEX-P*.

Governance Administration is done by ACONET, itself, located in Vienna.

Financing Funds come from the *Bundesminsteriums für Wissenschaft und Forschung*, the Austrian Ministry of Science and Research.

Connections EUnet and EARN.

b) UNA

Name *Universitäts-Netz Austria*, or Austrian University Network.

Date pre-1987

Purpose Connects Digital machines at Austrian Universities. It may eventually merge with ACONet.

Location Austria.

Configuration Connections between the Digital computers is accomplished via the PDN, *DATEX-P*, using X.25 protocols. Packet Network Systems (PSI) software is used to allow the use of the DECNET protocol. Typical DECNET network services are supported. Plans are underway to migrate to OSI protocols.

Financing Monies come from grants from *Bundesministeriums für Wissendeschaft und Forschung*, the Austrian Ministry of Science and Research.

Membership All universities in Austria

Connections There is a connection to EUnet. EARN and DFN connections are in progress.

A.2.10 Italy

a) INFNET

Name — *Instituto Nazionale Fisica Nucleare Network*, or the National Institute for Nuclear Physics Network.

Date — pre-1987

Purpose — Supports connections between sites of the Italian National Institute for Nuclear Physics.

Location — Italy.

Configuration — DECNET protocols are used.

Membership — There were 118 hosts as of July 1987.

Connections — SPAN.

A.2.11 Spain

a) Enet

Name — Enet is the Spanish division of EUnet. "E" stands for *España*.

Date — 1986

Location — Spain.

Configuration — The main protocol is UUCP. The UUCP *g* protocol is used over dial-up links, the *f* protocol over the Spanish X.25 PDN, *IBERPAC*, and the *t* protocol or SMTP over TCP/IP. Enet is linked internationally via *IBERPAC*.

Governance — The backbone machine is run by the *Departamento Ingenienría Telemática* (DIT), or the Department of Telematics Engineering at the *Universidad Politecnica de Madrid*.

Financing Local hosts pay their own transmission costs to the backbone directly. DIT expects to collect fees to recover costs incurred in running the backbone. Partial funding comes from the *Interconecion de recursos Informaticos* (IRIS), a national research plan for the coordination of Spanish research networks.

Membership There are 23 registered sites and 43 hosts. Sixty percent are commercial sites. The remainder are research sites.

Connections EARN and HEPnet.

b) RICA

RICA is a research network that connects universities in Andalusia. It uses Ean X.400 software and DECNET protocols or an X.25 network layer.

c) Ean in Spain

Date 1985

Purpose This is an Ean X.400 experimental research network.

Location Spain.

Configuration Ean X.400 software is run over the X.25 PDN *IBERPAC* with 9600bps and 1200bps connections.

Governance Administered and funded by IRIS.

Membership As of November, 1988, there were 22 hosts at 14 sites.

Connections There is a gateway to EUnet.

A.2.12 Greece

a) ARIADNE

Name Not available

Date 1989

Purpose	The objective of ARIADNE is to establish a widespread infrastructure for the academic and research and development community. OSI conformance is a prime goal.
Location	Greece
Configuration	The network connects 10 sites via a private X.25 packet-switched network. There is a commitment to migrate to OSI, with X.400 MHS, FTAM, and Network Management pilot projects underway.
Governance	The network is administered by the Greek National Research Centre, *Democritos*, located in Athens. The entire ARIADnet program is coordinated by the Greek representative to RARE and COSINE.
Financing	The network is funded by the Greek General Secretariat for Research and Technology and CEC.
Membership	There are currently over 100 hosts connected to ARIADNE.
Connections	ARIADNE provides Greece with IXI service, and has gateways to DARPA, EARN, and BITNET.

(Arvillas, 1990)

A.2.13 Yugoslavia

a) SIS

SIS, or the Social Information System, is the Yugoslavian PDN that may serve as a research network. Specialists hope to use it to offer mail, remote job entry, file transfer, and videotext.

A.3 Eastern Europe

Traditionally, networks in Eastern Europe have had few connections outside the Soviet Block due partly to the reluctance of both Communist governments to promote open communication and of the West to export and hardware software. The future status of networks in Eastern Europe, however, is extremely uncertain given the political upheavals that occurred in 1989 and continue to this day. The fall of Communist governments throughout Eastern Europe, and the recent reunification of Germany, will doubtless change the face of networking from what is outlined in this section.

A.3.1 International networks in Eastern Europe

a) IASnet

Name Institute for Automated Systems Network.

Date Under implementation.

Purpose IASnet is a network for Socialist countries. It provides institutions in Eastern Europe with access to Soviet databases and a few foreign databases.

Location Bulgaria, Hungary, former East Germany, Poland, Czechoslovakia, Cuba, Mongolia, and Vietnam.

Configuration The X.25 protocol connects the various informatics institutes. Connections outside IAS are done via an X.75 PDN in Austria called RADAUS, and an X.25 connection to the Finnish PDN, *Datapak*.

A.3.2 Former German Democratic Republic

Since the re-unification of East and West Germany, work is underway to provide DFN connections to institutions in the former German Democratic Republic. Among the main goals of the projects are:

- setting up a communications infrastructure for research institutes and organizations
- introduction and application of OSI-conformant services
- acquisition of experience in using the DFN services and in X.25 connectivity.

Initial connections will go to the Humboldt University in Berlin, the Institute for High-Energy Physics at Zeuthen, and the Technical University at Dresden (IES News, 1990c).

A.3.3 USSR

a) Academnet

Academnet connects the Institute for Automated Systems in Moscow to institutes in the Soviet Republics. There are connections to Leningrad, Riga, Kiev, Novosibirsk, and Dushanbe.

b) EARN in the USSR

The Soviet Academy of Science has recently been given permission to connect to EARN. The primary EARN node will reside at the Institute of Organic Chemistry in Moscow. The connection is expected to be operational by the end of 1990 (New Scientist, 1990).

A.3.4 Access in other Eastern European countries

There are connections to IASnet in Poland, the German Democratic Republic, Czechoslovakia, Hungary, and Bulgaria. Hungary also has an EUnet connection. There are no known networks in Romania or Albania.

A.4 Australasia

A.4.1 Australia

a) AARNet

Name The Australian Academic and Research Network.

Date AARNet supersedes SPEARNET and ACSnet, thus becoming Australia's main research network. It was installed and commissioned in May, 1990

Purpose The general purpose of AARNet is to support academic and research activities undertaken by higher education institutions and research organizations within Australia

Location The network spans the Australian continent.

Configuration TCP/IP and DECNET protocols. The use of X.25 and other protocols will be considered as future options. Internal links run at 48Kbps, the international link at 56Kbps. Services include all TCP/IP and DECNET applications including electronic mail, file transfer, news, and remote login.

Financing AARNet is funded by contributions from each participating institution, as well as 30 per cent funding from the Australian Research Council.

Governance The network is managed by the Australian Vice-Chancellor's Committee.

Membership

- Charles Stuart University
- Curtin University of Technology
- Deakin University
- Flinders University
- Griffith University
- James Cook University
- La Trobe University
- Macquarie University
- Monash University
- Murdoch University
- Northern Territory University
- Queensland University of Technology
- The Australia Defence Force Academy
- The Australia National University
- The Commonwealth Scientific and Industrial Research Organization
- The University College of Central Queensland
- The University College of Southern Queensland
- The University of Adeliade
- The University of Canberra
- The University of Melbourne
- The University of New England
- The University of New South Wales
- The University of Newcastle
- The University of Queensland
- The University of South Australia
- The University of Sydney
- The University of Tazmania
- The University of Technology, Sydney
- The University of Western Australia
- The University of Western Sydney
- The University of Wollongong
- Victorian University of Technology
- Western Australia College and Advanced Education

Connections The *Internet.*

(Huston, 1990b)

b) INFOPSI

INFOPSI is a DECNET network running over X.25 that connects about 20 research and academic sites in Australia and New Zealand. Mail and file transfer are supported.

b) CSIRONET

Name Commonwealth Scientific and Industrial Research Organization Network.

Date Development can be traced from 1962.

Purpose CSIRONET is a campus-level government research network.

Location Australia.

Configuration Inhouse protocols are used. Services are mail, remote login, remote job entry, and file transfer. Migration to X.25 for the network layer is underway. Virtual circuits are maintained between host machines at speeds of 48Kbps.

Financing Originally funded by the Australian Division of Computing Research which has since disbanded.

Membership There are eight academic institutions connected.

Connections Overseas connections are available through the PDNs, *MIDAS* and *TYMNET*.

A.4.2 New Zealand

a) DSIRnet

Name Department of Scientific and Industrial Research network.

Date 1977

Purpose Provides computer access to several large government installations.

Location New Zealand.

Configuration Inhouse DSIR protocols are used largely . An SNA emulation runs on top of DSIR protocols allowing remote login, remote job entry, and remote printing. An X.25 PAD allows DECNET access and mail. A few sites use exclusively DECNET and Coloured Book protocols.

Governance The network is the responsibility of the Government of New Zealand.

Financing Government of New Zealand.

A.5 South East Asia

A.5.1 International networks in South East Asia

a) AUSEAnet

Name Australasia and South East Asia Network.

Date 1986

Purpose AUSEAnet is a metanetwork supporting a Very Large Scale Integration (VLSI) project among the Association of South East Asian Nations countries (ASEAN) and Australia. The network allows the exchange of information about microelectronics techniques.

Location Thailand, Indonesia, Malaysia, Singapore, Brunei, the Philippines, and Australia.

Configuration UUCP and SUN-III protocols are used over X.25 networks. The UUCP portions are mainly outside of Australia. Each participating country will have its own network joined by a central machine in Indonesia.

Financing Provided by the Australian government.

Membership Currently at least nine ASEAN institutions are connected.

A.5.2 Malaysia

a) RangKoM

Name Rangkaian Komputer Malaysia.

Date 1987

Purpose The objective of the network is to link most universities in Malaysia together for network demonstration and development.

Location Malaysia.

Configuration Currently UUCP-based, however, there are plans to move toward X.25 and TCP/IP for long-haul connections. Services supported are mail, file transfer, and USENET news.

Governance Administered by a technical committee comprised of organizational representatives.

Financing Supported in part by a grant from the ASEAN-Australian Economic Cooperation Programme (AAECP), though this support was expected to end in 1989. Most operating expenses are born by the government of Malaysia. Sites not involved with AAECP pay their own costs.

Membership

- ASEAN Institute of Forest Management
- Malaysian Institute of Microelectronic Systems
- MARA Institute of Technology
- National University
- University of Agriculture
- University of Malaysia
- University of Science, Ipoh
- University of Science, Penhang
- University of Technology

Connections RangKoM is the Malaysian part of AUSEAnet.

A.5.3 Indonesia

a) UNInet

Development of this university research network began in 1988. It is sponsored by the Indonesian government in cooperation with research and commercial organizations as part of a five year plan. There are currently 12 sites registered. The intention of the plan is to link 45 government sponsored universities in Indonesia.

A.5.4 Thailand

a) TCSnet

Name Thai Computer Science Network.

Date 1988

Purpose TCSnet is a university network.

Location Thailand.

Configuration As in Australia's ACSnet, SUN-III protocols are used. All links are 2400bps dial-up. Services are mail, file transfer, remote job entry, and remote login.

Financing There is no committed funding. Presumably, the institutions pay for costs.

Membership
- Asian Institute of Technology
- Chulalongkorn University
- Prince of Songkhla University
- Thammasat University

Connections UUNET.

A.5.5 Access in other South-East Asian countries

There is a AUSEAnet connection in the Philippines. There are, however, no known systems in Brunei, Vietnam, Cambodia, Laos, and Myanmarr.

A.6 Southern Asia

A.6.1 India

a) ERNET

Name Educational and Research Network.

Date This network is not yet operational, but extensive planning has been done.

Purpose The purpose of this network is first to connect eight academic institutions in India, and secondly to promote communications technology.

Location India.

Configuration Protocols are to be OSI as much as possible. Because of the lack of a domestic PDN and noisy telephone lines, ERNET will use a satellite network as the national backbone running X.25. Packet radio is also expected to be used to overcome the same problem in metropolitan areas. Services are to include mail, file transfer, remote login, and database access.

Governance Administration is provided by a committee made up of personnel from the Department of Electronics.

Financing The Government of India.

Membership
- Department of Electronics, Government of India
- The Indian Institute of Science in Bangalore
- The Indian Institute of Technology at Delhi, Kampur Kharagpur, Madras, and Bombay
- The National Centre for Software Technology

Connections There are planned connections to JANET, UUNET, and CERN.

A.6.2 Access in other countries of Southern Asia

Sri Lanka has an AUSEAnet node and CGNET subscribers. Both Bangladesh and Nepal have CGNET nodes. Pakistan has no known networks.

A.7 Far East

A.7.1 International networks in the Far East

a) PACNET

PACNET, or the Pacific Network, is an informal group of organizations and computer hosts that serve as an academic network for the Pacific region. Countries connected are South Korea, Malaysia, Singapore, Indonesia, Australia, and Hong Kong. There is also one host in the U.S. Direct connections are to ACSnet, USENET, UUCP, EUnet, and JUNET. There are indirect connections to CDNnet, CSNET, and the *Internet*.

b) PACCOM

PACCOM, or Pacific Communications Network, is an in-progress plan to build a Pacific regional fiber optic internetwork. The network will provide access to a string of optical telescopes in Hawaii. Currently there are links from Hawaii (the hub) to the *Internet*. Future links are planned to Japan, Australia and New Zealand. TCP/IP and DECNET protocols will be used. The project is funded by NASA and the NSF.

A.7.2 Hong Kong

a) HARNET

Name Hong Kong Academic and Research Network.

Date 1986

Purpose The network connects higher education institutions in Hong Kong

Location Hong Kong.

Configuration Protocols used are UUCP (dial-up links), DECNET (leased lines), or X.25 (PDNs). Services supported are mail, file transfer, remote login, and USENET news.

Governance Each institution administers its own portion of the network.

Financing Funds are provided by each institution, although international calls are charged back to users.

Membership

- Baptist College
- Chinese University of Hong Kong
- City Polytechnic of Hong Kong
- Hong Kong Polytechnic
- University of Hong Kong

Connections There are links to Australia, the U.S., South Korea, and China.

A.7.3 South Korea

a) SDN

Name System Development Network.

Date 1982

Purpose The network's objective is to facilitate computer communications and resource sharing for research institutions in the Republic of Korea, and to act as a basis for network experimentation. It is a backbone network that connects the LANs of major institutions in Korea.

Location The Republic of Korea.

Configuration The basic protocol in use is TCP/IP, although UUCP is also supported. There are some UUCP links that run on top of TCP/IP that runs, in turn, over X.25. The Korean X.25 PDN, *DACOM*, is used for most links. International connections are also X.25. Speeds range from 2400bps to 9600bps. There are plans to convert to OSI protocols.

Governance Directed by the Korean Advanced Institute of Science and Technology.

Financing Sites pay for their own connections. International connects are charged on a per-use basis.

Connections UUCP and CSNET.

A.7.4 Japan

a) JUNET

Name Japan UNIX Network.

Date 1984

Purpose JUNET is a nationwide network that was created to promote exchange among researchers both inside and outside Japan. It also serves as testbed for network experimentation, most notably for the problem of Japanese character handling. It was patterned after USENET and CSNET in the U.S.

Location Extends throughout the major Japanese Islands with concentrations near Tokyo and Osaka.

Configuration The network is largely UUCP, with X.25 dial-up links. Some portions run UUCP over TCP/IP. There are no immediate plans to migrate to OSI protocols. Services include news and mail.

Governance Overall administration is the domain of the administrators of the individual backbone hosts.

Financing Domestic services are provided free of charge, but international communication costs are charged back to users.

Membership There are 55 organizations in the Tokyo area, 17 in the Osaka and Kyoto area, and a number of others spread around the main islands.

Connections UUCP, CSNET, ACSnet, SDN, BITNET, JANET and the *Internet*.

b) N1

Name Inter-university Computer Network.

Date 1981

Purpose N1 is the oldest academic network is Japan, but has been expanded into the *Science Information Network* by NACSIS (see A.7.4.c). It connects universities in Japan and provides access to special hardware and software. It was the first academic network in Japan.

Location Japan.

Configuration The system utilizes the domestic X.25 PDN, but runs higher level system-specific protocols over it to provide remote job entry, file transfer, mail, and news.

Governance There is no centralized organization.

Membership There 36 university computing centres connected to N1/NACSIS.
(Ishida, 1988)

c) NACSIS

Name National Center for Science Information Systems. This Center operates a network called the *Science Information Network*, but is commonly referred to as NACSIS. It is considered a successor to N1.

Date NACSIS began expanding N1 in 1987.

Purpose The network provides access to widely separated mainframes and supercomputers. It connects libraries, universities, and research institutes (Naito, 1989).

Location Concentrated on main Japanese Islands.

Configuration TCP/IP, X.25 and N1 protocols are all used in the network. Some links are converting to IP. The Japanese X.25 PDN forms the basis of most links, however, because of its high cost, high-speed links through leased lines are forthcoming. The N1 protocols support remote login, remote job entry, mail, and file transfer (Ishida, 1988).

Financing Purchase of the leased lines is supported by government grants.

Membership Seven university computing centres are connected.

A.7.5 Access in other Far East countries

There are BITNET and PACNET connections in Taiwan and an IASnet connection in Mongolia. China has no national networks, but there are links to CSNET and HARNET. There are no known networks in the Democratic Peoples Republic of Korea.

A.8 Latin America and the Caribbean

A.8.1 International networks in Latin American and the Caribbean

a) CARINET

Name Not available.

Date 1982

Purpose CARINET is a general purpose conferencing network for business and development organizations. There is a focus on developing regions, particularly Latin America. The purpose of the network is mainly technology transfer from industrially developed nations to those that are less developed.

Location It has nodes in Latin American and the Caribbean, as well as Africa, the Middle East, Asia, Europe, and North America.

Configuration The network is based on X.25 linked through domestic PDNs. It provides both interactive and batch communications, and a document ordering service.

Governance

CARINET is governed by a consortium consisting of:

- Agricultural Cooperative Development International
- The Daedalus Group
- Devres, Inc.
- Partnership of Productivity
- The Rodale Institute

Financing The network is user funded.

Membership Membership consists of local government and non-government agencies. There are currently about 250 users.

b) CATIENET

Name CATIENET is named after the *Centro Agronómico Tropical de Investigación y Enseñanza* (CATIE), or the Tropical Agricultural Research and Training Centre, a Central American/Caribbean based non-profit agricultural organization.

Date It was expected to be operational in 1989.

Purpose The goal of the network is to permit information exchange on agricultural and forestry topics. It provides access to databases and computer resources.

Location Costa Rica, Dominican Republic, El Salvador, Guatemala, Honduras, Nicaragua, and Panama.

Configuration The systems use TCP/IP over leased line. Current services are mail and file transfer; future services will include remote job execution and interactive database access.

Membership Membership include all members of CATIE, that is, all Spanish speaking countries in Central America and the Caribbean.

Connections Connections to BITNET are under consideration.

A.8.2 Mexico

a) UNAM

UNAM, is a network managed by *Universidad Nacional Autonomidad de México*, a national university in Mexico City. It consists of 15 sites connected by leased lines.

b) ITESM

Name Derived from *Istituto Tecnoló de Estudios Superiores De Monterrey*.

Date 1987

Purpose ITESM connects universities throughout Mexico.

Location Mexico.

Protocols. Links are by microwave and satellite. The main Mexico City campus network is a fiber optic Ethernet.

Governance. Administered by the ITESM Mexico City campus.

Membership. There are currently 26 ITESM campuses online.

Connections. Tentative plans exist to connect ITESM to NSFNET.

A.8.3 Access in other Latin American countries

Chile has access to BITNET and UUNET, as does Argentina. Haiti has BITNET access and Cuba has a IASnet node. There are no known networks, however, in Belize, Bolivia, Brazil, Columbia, Ecuador, French Guiana, Guyana, Paraguay, Peru, Surinam, Uruguay, and Venezuela.

A.9 Middle East

A.9.1 International networks in the Middle East

a) GulfNet

GulfNet is similar to EARN and BITNET with the same technologies and services. There are nine nodes in all, seven in Saudi Arabia and two in Kuwait.

A.9.2 Israel

a) ILAN

ILAN is the Israeli Academic Network and is a branch of EARN. Protocols were originally NJE, but are currently TCP/IP with plans to convert to OSI. Forty-seven hosts are connected country-wide.

A.9.3 Access in other Middle Eastern countries

Cyprus has access to EARN. There are only PDNs in Bahrain, Qatar, the United Arab Emirates, and Iraq. There are no networks known to be in Oman, Iran, Yemen, Jordan, Lebanon, and Syria.

A.10 Africa

A.10.1 Access in North African countries

While information about networks in this part of the world is limited, North African countries do have some network access. Egypt, for example, has a PDN, *ARENTO*, and a research network called ESTINET that may be eventually connected to EARN. There is also a network under development in Tunisia called *Afrimail*, an initiative designed to provide Ean mail services over any available media such as X.25 PDN services, leased lines, public telephone lines, and *Telex* links. Algeria and Morocco have EARN connections, but there are no networks in Libya.

A.10.2 Access in sub-Saharan Africa

There are no research networks in sub-Saharan Africa recognizable as similar to those in North America and Europe. This fact is due, in large part, to a definite lack of a land-based communications infrastructure to support computer connectivity. South Africa, however, is an

exception. There are internal networks in this country, but because of the international response to *Apartheid*, there are few outside connections. South Africa has applied for connection to EARN, BITNET, EUnet, UUNET, and CSNET, but has been turned down by all.

There is world-wide PDN-based conferencing system called CGNET that has connections in some sub-Saharan African countries. This system was founded by the Consultative Group on International Agricultural Research and is sponsored by a number of international agricultural research centres. It has 130 remote outposts worldwide. There are connections to CGNET in Mali, Burkina Faso, Niger, Ethiopia, and Zimbabwe.

There have also been discussions at the UN University Programme Advisory and Planning Meeting in April, 1989, about under taking a study about research networks with special emphasis on developing countries. One of a variety of activities singled out for action was participation in establishing ESANET (East and Southern African Network) that would link five universities in Kenya, Uganda, Tanzania, Zambia, and Zimbabwe (Wesley-Tanaskovic, 1990).

The remainder of network activity in this part of the world includes an EARN connection in the Ivory Coast and a packet-radio network (the use of radio to send packet-switched data) in Senegal called *Alternet*. There are no networks in Nigeria or Cameroon.

Index to Research networks

A.1 North America A-4
A.1.1 International networks in North America A-4
a) The *Internet* A-4
A.1.2 The United States A-5
a) NSFNET A-5
b) NSFNET mid-level networks A-6
CERFnet A-6
NYSERNet A-7
JVNCNet A-8
SURAnet A-8
OARnet A-9
CICNet A-10
Merit A-11
MIDnet A-11
NCSAnet A-12
MRNet A-13
WESTNET A-14
NorthWestNet A-14
BARRNet A-15
THEnet A-16
Sesquinet A-16
USAN A-17
e) MFEnet A-18
f) SPAN A-18
A.1.2.1 Other important U.S. networks A-19
a) BITNET A-19
b) USENET A-19
c) UUCP A-20
d) HEPnet A-21
A.1.3 Canada A-21
a) NetNorth A-21
b) CDNnet A-22
c) CAnet A-23
d) CAnet mid-level networks A-24
NLnet A-24
NSTN A-24
AccessNB A-25
The Prince Edward Island Network A-26
RISQ A-26
Onet A-27
MBnet A-28

Sask#net A-29
ARnet A-29
BCnet A-30
e) ISTC Network A-31

A.2 Europe A-32
A.2.1 International networks in Europe A-32
a) IXI A-32
b) EUnet A-33
c) EARN A-34
A.2.2 United Kingdom A-35
a) JANET A-35
b) UKnet A-36
A.2.3 Ireland A-37
a) HEANET A-37
A.2.4 Scandinavia A-37
a) NORDUnet A-37
b) UNINETT A-38
c) FUNET A-39
d) DENet A-39
e) SUNET A-40
f) EUnet in Iceland A-41
A.2.5 Germany (Formerly West Germany) A-41
a) DFN (WIN) A-41
b) AGFnet A-42
c) Dnet A-43
A.2.6 France A-43
a) REUNIR A-43
b) ARISTOTE A-44
c) FNET A-45
A.2.7 The Netherlands A-45
a) SURFnet A-45
A.2.8 Switzerland A-46
a) SWITCH A-46
b) CERN A-46
A.2.9 Austria A-46
a) ACONet A-46
b) UNA A-47
A.2.10 Italy A-48
a) INFNET A-48
A.2.11 Spain A-48
a) Enet A-48
b) RICA A-49

A.2.12 Greece A-49
a) ARIADNE A-49
A.2.13 Yugoslavia A-50
a) SIS A-50

A.3 Eastern Europe A-50
A.3.1 International networks in Eastern Europe A-51
a) IASnet A-51
A.3.2 Former German Democratic Republic A-51
A.3.3 USSR A-51
a) Academnet A-51
b) EARN in the USSR A-52
A.3.4 Access in other Eastern European countries A-52

A.4 Australasia A-52
A.4.1 Australia A-52
a) AARNet A-52
b) INFOPSI A-53
b) CSIRONET A-54
A.4.2 New Zealand A-54
a) DSIRnet A-54

A.5 South East Asia A-55
A.5.1 International networks in South East Asia A-55
a) AUSEAnet A-55
A.5.2 Malaysia A-55
a) RangKoM A-55
A.5.3 Indonesia A-56
a) UNInet A-56
A.5.4 Thailand A-56
a) TCSnet A-56
A.5.5 Access in other South-East Asian countries A-57

A.6 Southern Asia A-57
A.6.1 India A-57
a) ERNET A-57
A.6.2 Access in other countries of Southern Asia A-58

A.7 Far East A-58
A.7.1 International networks in the Far East A-58
a) PACNET A-58
b) PACCOM A-59
A.7.2 Hong Kong A-59
a) HARNET A-59

A.7.3 South Korea A-59
a) SDN A-59
A.7.4 Japan A-60
a) JUNET A-60
b) N1 A-61
c) NACSIS A-61
A.7.5 Access in other Far East countries A-62

A.8 Latin America and the Caribbean A-62
A.8.1 International networks in Latin American and the Caribbean A-62
a) CARINET A-62
b) CATIENET A-63
A.8.2 Mexico A-64
a) UNAM A-64
b) ITESM A-64
A.8.3 Access in other Latin American countries A-64

A.9 Middle East A-65
A.9.1 International networks in the Middle East A-65
a) GulfNet A-65
A.9.2 Israel A-65
a) ILAN A-65
A.9.3 Access in other Middle Eastern countries A-65

A.10 Africa A-65
A.10.1 Access in North African countries A-65
A.10.2 Access in sub-Saharan Africa A-65